Proceedings of Congress Papers

9th Iranian Joint Congress on Fuzzy and Intelligent Systems

(CFIS2022), March 2-4, 2022

Higher Education Complex of Bam, Iran

20th Iranian Conference on Fuzzy Systems (ICFS2022)

18th Conference on Intelligent Systems (CIS2022)

5th Swarm Intelligence and Evolutionary Computation Conference (CSIEC2022)

Proceedings of 9th Iranian Joint Congress on Fuzzy and Intelligent Systems
Edited by: Fatemeh Barani, Me’raj Abdi
Higher Education Complex of Bam
April, 2022

In the Memory of L. A. Zadeh

The father of fuzzy

(1921-2017)

In the memory of Dr. Bahram Sadeghpour Gildeh

Professor, Department of Statistics, Ferdowsi University of Mashhad

(1966-2021)

Time-table
9th Joint Congress on Fuzzy and Intelligent Systems
(First day)
Iran Daylight Time (GMT +03:30)

HALL 3	HALL 2	HALL 1	Wednesday March/2/2022
Opening Ceremony			08:00-10:00
Tea Break			10:00-10:30
Keynote Speech: **Evolutionary Intelligence in (Big) Data Analytics and Optimization** **Prof. Amir H. Gandomi, University of Technology Sydney, Australia** Chairman: Prof. Mashaallah Mashinchi			10:30-11:30
Lunch Break			11:30-13:00
General Meeting of the Iranian Fuzzy Systems Society			13:00-14:00
Session: **Text Processing and Social Networks**	**Special Session:** **AI for Good**	**Special Session:** **Fuzzy Optimization 1**	14:00-16:00
Tea Break			16:00-16:15
Keynote Speech: **Overview of new families of fuzzy implications and their applications in fuzzy systems** **Prof. Michal Baczynski, University of Silesia in Katowice, Katowice, Poland** Chairman: Prof. Seyyed Naser Hosseini			16:15-17:15
Tea Break			17:15-17:30
Session: **Swarm Intelligence and Evolutionary Computation**	**Special Session:** **(Fuzzy) Logical algebraic structures 1**	**Special Session:** **Fuzzy Numerical Analysis**	17:30-19:30
Keynote Speech: **The Sound of Health** **Prof. Mohammad-R Akbarzadeh-T, Ferdowsi University of Mashhad (FUM), Iran** Chairman: Prof. Kambiz Badie			19:30-20:30

Time-table
9th Joint Congress on Fuzzy and Intelligent Systems
(Second day)
Iran Daylight Time (GMT +03:30)

HALL 3	HALL 2	HALL 1		Thursday March/3/2022
Session: Statistics and Probability in Fuzzy	**Session: Intelligent Systems and Applications 2**	**Special Session: Fuzzy automata and graphs**	08:00-10:20	08:00-10:00
Tea Break				10:00-10:30
Tea Break				
Keynote Speech: **Role of Intelligent systems in the Interpretation of Deep Neural Networks in a Post-pandemic World** **Prof. Saman K. Halgamuge, University of Melbourne, Australia** Chairman: Prof. Mehdi Eftekhari				10:30-11:30
Lunch Break				11:30-13:00
General Meeting of the Iranian Intelligent Systems of Scientific Society		**Special Session: Copula Theory and Its Applications**	13:00-16:00	13:00-14:00
Session: Data Mining	**Session: Applications of optimization and operation research in fuzzy**			14:00-16:00
Tea Break				16:00-16:15
Keynote Speech: **Data-Driven Fuzzy Modeling** **Prof. Irina Perfilieva, University of Ostrava, Czech Republic** Chairman: Prof. Amir Daneshgar				16:15-17:15
Tea Break				17:15-17:30
	Session: Machine Learning	**Special Session: (Fuzzy) Logical algebraic structures 2**		17:30-19:30
Keynote Speech: **Some recent extensions of fuzzy integrals applied to the computational brain problem** **Prof. Humberto Bustince, Universidad Publica de Navarra, Spain** Chairman: Prof. Esfandiar Eslami				19:30-20:30

Time-table
9th Joint Congress on Fuzzy and Intelligent Systems
(Third day)
Iran Daylight Time (GMT +03:30)

HALL 3	HALL 2	HALL 1	Friday March/4/2022
	Session: Bioinformatics	**Special Session: Fuzzy Optimization 2**	08:00-10:00
Tea Break			10:00-10:30
Keynote Speech: **Possibility-Based Perception: A Reflection on Recognition of Shape Pattern** **Prof. Kambiz Badie, University of Tehran & ITRC, Iran** Chairman: Prof. Mohammad-R Akbarzadeh-T			10:30-11:30
Lunch Break			11:30-13:00
Session: Machine Vision and Image Processing	**Session: Fuzzy Systems and Applications**	**Special Session: (Fuzzy) Logical algebraic structures 3**	13:00-15:00
Tea Break			15:00-15:15
Keynote Speech: **Derivative of fuzzy functions and its application in fuzzy differential equations** **Dr. Alireza Khastan, Department of Mathematics, IASBS, Iran** Chairman: Prof. Reza Saadati			15:15-16:15
Tea Break			16:15-16:30
	Session: Intelligent Systems and Applications 2	**Session: Fuzzy Mathematical and Numerical Analysis**	16:30-18:30
Tea Break			18:30-18:45
Closing Ceremony			18:45-20:45

Message from the congress chair

In the name of God, the compassionate and the merciful Greetings and courtesy to all dear scholars, guests, and esteemed speakers.

I sincerely thank all for accepting our invitation and adding to the richness of this conference with their presence.

I would like to thank the Iranian Fuzzy Systems Association and the Iranian Intelligent Systems Association, which awarded the hosting privilege of this congress to the Higher Education Complex of Bam and continued supporting the 9th Joint Iranian Fuzzy and Intelligent Systems Congress.

I would also like to thank the Vice-Chancellor for Research of the Shahid Bahonar University of Kerman, scientific associations, and other esteemed contributors that have accompanied us in organizing this event.

I am grateful to my dear colleagues and professors who worked round the clock in scientific and executive committees, especially the respected scientific secretaries of the congress, Prof. Mashaallah Mashinchi and Prof. fereshteh Forouzesh, and the executive secretary of the congress, Dr. Me'raj Abdi, which I have witnessed the efforts and concerns of these colleagues in organizing this congress in recent months.

I thank God for giving me the opportunity of being with my dear colleagues in this congress. I am very sorry that the complex was deprived of the pleasure of talking and visiting all participants in person due to the Coronavirus pandemic. However, I am still proud that the Higher Education Complex of Bam hosts virtually great scholars like you. We appreciate this opportunity for further development and progress of our university.

Last year, we lost great scientists and professors, and I consider it necessary to cherish the memory of these late myths of the country: the late professor Bahram Sadeghpour Gildeh, the late professor Mahmoud Lashkarizadeh Bami, and others who said goodbye to the mortal world this year and joined immortality. The memory of these scholars will never vanish from our hearts, and their names will never disappear from the scope of science and knowledge of this region.

I would like to inform you that the Higher Education Complex of Bam, which is more than thirty years old, has always been active in promoting science and research in society. This complex successfully hosted the 12th Iranian Intelligent Systems Conference in 2013 in collaboration with the Iranian Intelligent Systems Association, the Institute of Electrical and Electronics Engineers, and the Department of Electrical Engineering of Shahid Bahonar University of Kerman. The event was a venue for convening researchers and experts as well as disseminating new findings in the field of smart systems and their applications in society.

In addition, with the contribution of the cooperation of Shahid Bahonar University of Kerman and Graduate University of Advanced Technology, this university hosted the first and third conferences in Evolutionary Computing and Collective Intelligence in 2015 and 2017.

This year we were honored to host the 9th Congress of Intelligent and Fuzzy Systems in cooperation with scientific associations and universities in Kerman province, the 20th Conference on Fuzzy Systems, the 18th Conference on Intelligent Systems, and the 5th Conference on Evolutionary Computing and Collective Intelligence. Certainly, this congress is a great opportunity to share the views and opinions of experts and technicians in fuzzy science and its application in various sciences.

Finally, I would like to express my gratitude to all who, with their financial and spiritual support, have helped Higher Education Complex of Bam to host this congress in the most glorious way possible. I am also grateful to the board of governors of the complex, the executive and the scientific committees, the panel of reviewers, the esteemed faculty members, the administrative staff of the complex, and those who contributed to organizing and hosting this conference. I am utterly thankful to the esteemed professors of the universities and students for their active presence and participation in the congress, and to the researchers, experts, and dignitaries who added to the scientific richness of the discussions by sharing an article, delivering a lecture, or participating in the talks.

I hope that, despite Covid-19 limitations, this congress was a good place to present the scientific achievements of researchers in various fields.

Mohammad Ali Nourollahi
Chair of the congress
Dean of the Higher Education Complex of Bam

Message from the president of Iranian Fuzzy Systems Society

First of all, I take the holding of the 9th Joint Congress on Fuzzy and Intelligent Systems in Higher Education Complex of Bam as a good opportunity, and I express my gratitude to all of my dear colleagues who have helped in holding this congress. My deepest gratitude goes to my hard-working colleagues and managers and also the esteemed dean of the Higher Education Complex of Bam that despite the limitations caused by the COVID-19 pandemic and even financial difficulties, they tried their best to hold this congress in the best possible way; however, I consider it necessary to mention the following points:

Based on the Scopus Scientific Database, in the last ten years, the Islamic Republic of Iran has been ranked in third place in the production of the papers in the field of fuzzy systems after China and India, while the United States is ranked in fourth place. Although our country has 1% of the world's population, it has managed to account for 13% of the production of papers in the field of fuzzy systems in 2021, while the production of documents in all scientific fields has been 2%. Therefore, the role of fuzzy systems theory in the production of science and scientific papers has been very effective in promoting the scientific position of our country in the region and in the world.

It is noteworthy that despite the growth of 50% of Humanities Research Papers in the world during the last 5 years, in our country it has grown more than 200%, and also we have had the increasing growth of up to 100% in the use of the theory of fuzzy systems in materials science, up to 50% in agriculture and biotechnology, and up to 20% in pharmacy. Therefore, the above information shows that the utilization of the theory of fuzzy systems among scientists and researchers in our country is increasing, and this knowledge has been used in new fields such as intelligent systems, artificial intelligence, humanities, basic sciences and engineering sciences.

Perhaps the day Professor Lotfizadeh introduced fuzzy logic, he did not believe that this theory, after nearly six decades, was still one of the fascinating research topics in the world, which fascinated many researchers. After all these years, what still makes fuzzy logic and intelligent systems popular scientific research, has to do with its integration with the current needs of society, the human mind and also its applicability in various fields related to human life; however, we should acknowledge that our country's researchers have paid less attention to its practical aspects. Therefore, science and technology systems should be planned so that applied research can be provided and the needs of society can be met regarding the application of fuzzy and intelligent systems.

On this basis, it is suggested that based on the scientific-elite demand in this field, a scientific association called

"Science and Technology of Fuzzy and Intelligent Systems Association"

be established under the vice-presidency's supervision for Science and Technology. Therefore, as the President of the Iranian Fuzzy Systems Society, I ask all my dear colleagues and researchers who work in this field to help and guide us to provide the necessary conditions for the establishment of the association in question.

I wish for the scientific and technological self-sufficiency of our country, and thanks again to those who were involved in holding this scientific event.

Mohammad Mehdi Zahedi
The President of the Iranian Fuzzy Systems Society

Message from the Vice President of the Intelligent Systems Scientific Society of Iran

Nowadays, the modeling and the analysis of complex phenomena, require new scientific and accurate methods, and one can no longer be satisfied with the scientific methods that seek only analytical and purely accurate answers, since the analytical methods in the past have failed to solve many issues. Soft computing can be related to the new computational method in mathematics, computer science, machine learning, artificial intelligence and many other applied fields. These methods make it possible to deal with the shortcomings of analytical methods by modeling and studying the very complex phenomena and to address the shortcoming regarding the obtained results.

Nowadays, the current knowledge indicates the increasing commonality of different sciences. Therefore, in the discussion of soft computing, the close connection between mathematics, statistics, computer, electrical engineering, humanities and other related fields cannot be denied. It can be said that the more these branches of science and technology are connected, the more attractive they become, and it is no longer possible to create a strong barrier between them in the traditional way, and the old barrier must also be removed little by little. Anyone should take step in the borders of the field of science and engineering, and this is certainly the case with fuzzy and intelligent systems. Today we see that mathematics and fuzzy logic have found many applications in the humanities such as psychology and sociology and management, which are mostly inaccurate concepts.

In recent years, the Intelligent Systems Society has achieved significant success in this regard with the aim of expanding and advancing the scientific development of intelligent systems and the qualitative development of specialized forces and improving education and research in related fields, including the synergy of sciences related to this field through holding numerous national and international conferences, publishing new books in the field of uncertainty theory and effective communication with the pioneers and innovators of this science all over the world.

At the end, it should be mentioned that the Intelligent Systems Society, is pleased to support the 9th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS2022) which includes three successful periodic conferences: the 20th Iranian Fuzzy System Conference (ICFS2022), 18th Iranian Intelligent Systems Conference (CIS2022), and the 5th Conference on Swarm Intelligence and Evolutionary Computation (CSIEC2022), and it is hoped that holding this Congress, and holding it in person after the end of the Coivid-19 Pandemic, will attract the attention of all professors and researchers in the fields related to the mentioned topics and will cause the development of interdisciplinary subjects, to provide a suitable environment for scientific growth, discussion and exchange.

With respect
Behrouz Fathi Vice President of the Intelligent Systems Scientific Society of Iran

Message from the Scientific Secretaries of Congress

We thank God for helping us to hold the 9th Joint Congress of Iranian Fuzzy and Intelligent Systems from March 2 to 3, 2022, in cooperation with the Iranian Fuzzy Systems Association, the Iranian Intelligent Systems Association, the International Fuzzy Association, and many scientific associations related to the scopes of the congress.

We are also thankful to colleagues, students, staff, and the support of esteemed officials of the Higher Education Complex of Bam.

It is hoped that the result of the organizers' efforts will provide the satisfaction of the professors and students, and as the result, with their kindness and care, they will overlook the possible shortcomings.

Planning for the congress began in September 2020, and we faced a considerable welcome of professors, students, and researchers in the country. However, due to time constraints, we had to select a number of research articles of higher quality to present at the congress.

Applicants for the conference submitted about 260 research articles to the conference secretariat, which after a two-phase review (initial and specialized review) by at least 2 reviewers, 120 articles were accepted in the form of lectures and 20 articles in the form of posters. For each accepted lecture, 15 minutes of presentation time and 5 minutes of question and answer time were considered. The presentations were received from the authors in recorded format due to the virtual module of the congress.

In addition, eight prominent professors from inside and outside the country were invited for keynote speeches, which were presented virtually. The duration of each Keynote speech was 45 minutes along with 15 minutes of questions and answers.

There were also 5 workshops which were held virtually for 3 to 4 hours. The annual general assembly of the two associations of Fuzzy Systems and Intelligent Systems was also held alongside the congress.

Organizing this congress is the result of the cooperation of many individuals, especially the tireless efforts of the executive secretary of this congress. We would like to thank the board of governors of the Higher Education Complex of Bam, the Iranian Fuzzy Systems Association, and the Iranian Intelligent Systems Association, other supporters whose names are mentioned in this booklet.

Finally, we would like to appreciate the respected members of the Scientific Committee of the Congress, especially Prof. Arsham Boroumand Saeid and Associate Prof. Marjan Kouchaki Rafsanjani, who accompanied us in reviewing the received manuscripts round the clock.

Best regards,
Scientific Secretaries of Congress
Mashaallah Mashinchi
Fereshteh Forouzeh

Message from the Secretary of the Executive Committee of Congress

Praise be to the Glorious God who gave us existence and guided us and made us proud to be the companions of science and knowledge so that we can be blessed with this infinite ocean.

Whosoever served the creature the world will be eternalized his nature

After successfully holding the 12th Iranian Conference on Intelligent systems in 2014, the First and the Second Swarm Intelligence and Evolutionary Computation in 2016 and 2018, The Higher Education Complex of Bam is going to host the 9th Joint Congress on Iranian Fuzzy and Intelligent Systems (CFIS2022) with the support of the Iranian Fuzzy Systems Society and the Intelligent Systems of Scientific Society **on March 2-4, 2022**. Congress CFIS2022 includes three successful periodic conferences: the 20th Iranian Fuzzy System Conference (ICFS2022), the 18th Iranian Intelligent Systems Conference (CIS2022), and the 5th Conference on Swarm Intelligence and Evolutionary Computation (CSIEC2022).

With all due respect and honor, I would like to express my gratitude to the esteemed faculty members, students and researchers in the field of fuzzy and intelligent systems who added richness to this Congress by presenting their valuable papers.

Being supported by the Iranian Fuzzy Systems Society and the Intelligent Systems Scientific Society of Iran and being obtained a license in 2021, the executive affairs of the Congress began their activities from the beginning of Fall 2021. In the wake of the COVID-19 pandemic, and based on the decision of the Policy Council, the virtual nature of the Congress was put on the agenda of the Executive Committee, and accordingly, all matters, including site preparation for receiving papers and other matters, were done according to schedule. Sixty-two members of the Scientific Committee and about 100 reviewers with various specializations from all over the country and the world participated in explaining the general policies of the Congress and their implementation. In order to make this Congress more glorious and practical, 8 keynote speeches and 5 specialized workshops were selected to be presented in the Congress.

Here, I would like to thank the esteemed President of the Higher Education Complex of Bam as the President of the Congress, the esteemed Vice President for Education and Research, as well as the esteemed Director of Research Affairs, who always supported us in holding the Congress. I also express my gratitude for the financial and spiritual support of the Vice Chancellor for Research of the Shahid Bahonar University of Kerman and other supporters of the Congress, who supported us for the successful implementation of the Congress.

It is gratifying that I had a very valuable experience with the esteemed members of the Policy Council of the Congress, especially the kind and honorable professor, Prof. Mashaallah Mashinchi as the chair of the program committee, who transferred his many years of experience nationally and international scientific research fields and did not hesitate to convey it to the executive committee of Congress. I would also like to take this opportunity to thank Prof. Fereshteh Foruzesh as the Scientific Secretary of the Congress, Prof. Arsham Boroumand Saeid and Prof. Marjan Kouchaki Rafsanjani who have always been our assistants in decision-making over the past year, especially in managing and reviewing the papers.

It is worth mentioning that 260 paper codes were received on the Congress website, and after the initial review of the papers by at least two expert reviewers, 120 papers were selected as lectures, and 20 papers as posters.

I would like to express my sincere gratitude to all of the compassionate and hard-working colleagues of the Executive Committee of the Congress, including Ms. Mina Mirhosseini, Ms. Fatemeh Barani, Dr. Tayyebeh Askari Javaran, Dr. Mohadeseh Soleimanpour, Ms. Mahdieh Mozafari, Dr. Seyyed Hossein Mahdavi, Ms. Anahita Amirshojai, Dr. Zeinab Khatoon Pourtaheri, Ms. Mahta Bedrood, Mr. Abdul-Hamid Bahroloom, Mr. Hassan Hadidi, Ms. Nafiseh Hadidi and I humbly appreciate and thank other colleagues, visiting professors and students who have been our partners and companions in holding this glorious event.

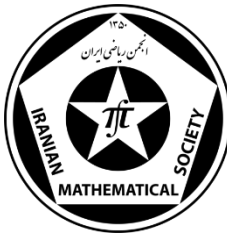
It is worth mentioning that the ISC index license has been obtained with the cooperation and efforts of Dr. Bijan Emamipour, the esteemed director of Research Affairs, and the IEEE index license has also been the result of the efforts of Ms. Fatemeh Barani. I hereby appreciate their efforts.

I have to honor the memory of the late Professor, Dr. Bahram Sadeghpour Gildeh, for his valuable and lasting services, who always was a moral character for many students and me.

It is hoped that the round-the-clock efforts of the members of the Executive Committee will satisfy the esteemed participants and supporters of the Congress, and it is also hoped that we will take a short step in promoting the scientific research objectives of our country.

With Best Regards
Me'raj Abdi
The Secretary of the Executive Committee of Congress

Scientific and Industrial Supporters of Congress



Members of Program Committee

Name	Institute
Prof. Saeid Abbasbandy	*Imam Khomeini International University, Iran*
Dr. Hamed Ahmadzade	*University of Sistan and Baluchestan, Iran*
Prof. Mohammad-R. Akbarzadeh-T	*Ferdowsi University of Mashhad, Iran*
Prof. Tofigh Allahviranloo	*Bahcesehir university, Istanbul, Turkey*
Prof. Mohammad Amini	*Ferdowsi University of Mashhad, Iran*
Dr. Mohsen Arefi	*Birjand University, Iran*
Dr. Alireza Askarzadeh	Graduate University of Advanced Technology, Kerman, Iran
Prof. Michal Baczynski	University of Silesia in Katowice, Poland
Prof. Kambiz Badie	*University of Tehran; Iran Telecommunication Research Center, Iran*
Dr. Mahmood Bakhshi	*University of Bojnord, Iran*
Prof. Arsham Borumand Saeid	*Shahid Bahonar University of Kerman, Iran*
Prof. R. A. Borzooei	*Shahid Beheshti University, Iran*
Prof. Amir Daneshgar	*Sharif University of Technology, Iran*
Prof. Bernard De Baets	Ghent University, Belgium
Prof. Vali Derhami	*University of Yazd, Iran*
Prof. Antonio Di Nola	University of Salerno, Italy
Dr. Sadegh Dorri Nogoorani	Tarbiat Modares University, Iran
Dr. Mahdi Eftekhari	*Shahid Bahonar University of Kerman, Iran*
Prof. Esfandiar Eslami	*Shahid Bahonar University of Kerman, Iran*
Prof. Behrouz Fathi-Vajargah	Gilan University, Iran
Dr. Mahmood Fazlali	Shahid Beheshti University, Iran
Dr. Fereshteh Forouzesh	Higher Education Complex of Bam, Iran
Prof. Seyed Naser Hosseini	*Shahid Bahonar University of Kerman, Iran*
Prof. Gunther Jaeger	University of Applied Sciences Stralsund, German
Dr. Mahdi Jmpour	*Quchan University of Technology, Iran*
Prof. Mohammad Jamshidi	The University of Texas, San Antonio, USA
Prof. Ehsanollah Kabir	Tarbiat Modares University, Iran
Prof. Ihsan Kaya	Industrial Engineering Yıldız Technical University, Turkey
Dr. Farshid Keynia	Graduate University of Advanced Technology, Kerman, Iran

Dr. Alireza Khastan	Institute for Advanced Studies in Basic Sciences, Zanjan, Iran
Dr. Marjan Kuchaki Rafsanjani	Shahid Bahonar University of Kerman, Iran
Dr. Fatemeh Kouchakinezhad	Iranian Journal of Fuzzy Systems, Iran
Prof. Weldon Lodwick	University of Colorado Denver, USA
Pro. Nezam Mahdavi-Amiri	Sharih University of Technology, Iran
Prof. M. Mashinchi	*Shahid Bahonar University of Kerman, Iran*
Dr. Mohamad Reza Mashinchi	Payame-Noor University - Kerman, Iran
Prof. Radko Mesiar	Slovak University of Technology in Bratislava, Slovak
Prof. Mohammad Bagher Menhaj	Amirkabir University of Technology, Iran
Prof. Hassan Mishmast Nehi	*University of Sistan and Baluchestan, Iran*
Prof. Gholam Reza Mohtashami Bordazran	*Ferdowsi University of Mashhad, Iran*
Prof. Susana Montes	University of Oviedo, Spain
Prof. John Mordeson	Creighton University, USA
Dr. S. H. Nasseri	University of Mazandaran, Iran
Prof. Hossein Nezamabadi-pour	*Shahid Bahonar University of Kerman, Iran*
Dr. Abbas Parchami	*Shahid Bahonar University of Kerman, Iran*
Dr. Mir Mohsen Pedram	Kharazmi University, Iran
Dr. Akbar Pad	*University of Bojnord, Iran*
Prof. Dan Ralescu	University of Cincinnati, United States
Dr. Reza Saadati	Iran University of Science and Technology, Iran
Prof. Nasser Sadati	Sharif University of Technology, Iran
Prof. Arash Sharifi	Science and Research Branch, Islamic Azad University, Iran
Dr. Ayyub Sheikhi	Shahid Bahonar University of Kerman, Iran
Prof. Fu-Gui Shi	Beijing Institute of Technology, China
Prof. Alexander Sostak	University of Latvua, Latvia
Prof. Seyed Mahmoud Taheri	*University of Tehran, Iran*
Prof. Mohammad Teshnehlab	*Khajeh Nasir* al-Din *Toosi University of Technology, Iran*
Prof. Ali Torabi	*University of Tehran, Iran*
Prof. Ali Vahidian Kamyad	Ferdowsi University of Mashhad, Iran
Prof, Jose Luis Verdegay	University of Granada, Spain
Dr. Zeshui Xu	Sichuan University, China
Dr. Mohammad Ali Yaghoobi	*Shahid Bahonar University of Kerman, Iran*
Prof. M. M. Zahedi	*Graduate University of Advanced Technology, Kerman, Iran*

Members of Executive Staff

General Chair	Dr. Mohammad Ali Noorolahi
Chairs of Program Committee	Prof. Mashaallah Mashinchi Dr. Fereshteh Forouzesh
Chair of Executive Committee	Dr. Me'raj Abdi
IT Committee Members	Mina Mirhosseini: Chair Dr. Tayebeh Askari Javaran Dr. Mohadese Soleimanpour-moghadam Abdolhamid Bahrololoum
Registeration Committee Members	Dr. Tayebeh Askari Javara: Chair Dr. Zeinab Khatoun Pourtaheri Mr. Hasan Hadidi
Secretariat Members	Dr. Seyyed Hossein Mahdavi Mahta Bedrood
Public Relations Committee	Dr. Mohammad Mahmoodi Rad Mojtaba Sadeghian Mahdiye Farzannia
Publishing Committee Members	Fatemeh Barani: Chair Mina Mirhosseini Nafiseh Hadidi Dr. Tayebeh Askari Javaran Dr. Bijan EmamipourDr. Mohadese Soleimanpour-moghadam Abdolhamid Bahrololoum Dr. Zeinab Khatoun Pourtaheri
Finance and Support Committee	Dr. Hamzeh Dehghani: Chair Support: Hossein Azimi Finance: Hasan Azimi
Industrial Communication and Workshop Committee Members	Dr. Hadi Farzan Salimeh Ghanbari Dr. Maryam ZiaAbadi Mohammad Hossein Forouzanfar
International Relations Committee Members	Anahita Amirshojaei: Chair

Members of Congress Policy Council

Name	Institute
Dr. Me'raj Abdi	Higher Education Complex of Bam
Prof. Saeid Abbasbandy	Imam Khomeini International University
Prof. Mohammad-R. Akbarzadeh-T	Ferdowsi University of Mashhad
Dr. Arsham Borumand Saeid	Shahid Bahonar University of Kerman
Prof. Esfandiar Eslami	Shahid Bahonar University of Kerman
Prof. Behrouz Fathi-Vajargah	Gilan University
Dr. Fereshteh Forouzesh	Higher Education Complex of Bam
Prof. Mashaallah Mashinchi	Shahid Bahonar University of Kerman
Prof. Hossein Nezamabadi-pour	Shahid Bahonar University of Kerman
Dr. MohammadAli Noorollahi	Higher Education Complex of Bam
Prof. Seyed Mahmoud Taheri	University of Tehran

The Referees Committee of CFIS2022

Name	Institute
Saeed Abbasbandy	Imam Khomeini International University
Me'raj Abdi	Higher Education Complex of Bam
Amir Abdollahi	Shahid Bahonar University of Kerman
Khadijeh Abolpour	Islamic Azad University
Mahboubeh Afzali	Kerman Graduate University of Technology
Hamed Ahmadzadeh	University of Sistan and Baluchestan
Mohammad Alaei	Vali-e-Asr University of Rafsanjan
Laaya Aliahmadi	Shahid Bahonar University of Kerman
Mehdi Allahdadi	University of Sistan and Baluchestan
Tofigh Allahviranloo	Bahcesehir International University Turkey
Mohammad Amini	Ferdowsi University Of Mashhad
Alireza Arabpour	Shahid Bahonar University of Kerman
Tayyebeh Askari Javaran	Higher Education Complex of Bam
Mahmood Bakhshi	University of Bojnurd
Fatemeh Barani	Higher Education Complex of Bam
Mojtaba Barkhordari	Shahid Bahonar University of Kerman
Arsham borumand Saeid	Shahid Bahonar University of Kerman
Behrang Chaboki	Vali-e-Asr University of Rafsanjan
Mohammad Bagher Dowlatshahi	Lorestan University
Sepehr Ebrahimi	University of Yazd
Mehdi Eftekhari	Shahid Bahonar University of Kerman
Mojtaba Eftekhari	Shahid Bahonar University of Kerman
Nasibeh Emami	Kosar University of Bojnord
Azadeh Alsadat Emrani Zarandi	Shahid Bahonar University of Kerman
Sadegh Eskandari	University of Guilan
Esfandiar Eslami	Shahid Bahonar University of Kerman
Fakhrosadat Fanian	Shahid Bahonar University of Kerman
Hamed Farahani	Chabahar University of Maritime Sciences
Farideh Farsad	Higher Education Complex of Bam
Hadi Farzan	Higher Education Complex of Bam
Hamideh Fatemidokht	Shahid Bahonar University of Kerman
Behrouz Fathi-Vajargah	University of Guilan
Fereshteh Forouzesh	Higher Education Complex of Bam
Mohammad Ghasemzadeh	University of Yazd
Shokoufeh Ghorbani	Shahid Bahonar University of Kerman
Mahdieh Ghazvini Kor	Shahid Bahonar University of Kerman
Farhad Hamidi	University of Sistan and Baluchestan
Mohammad Hamidi	Payam-noor University of Kashan
Hamzeh Harfsheno	Farhangian University of Kashan
Mahsa Hasanshahi	Shahid Bahonar University of Kerman

Soodeh Hosseini	Shahid Bahonar University of Kerman
Mohammad Izadi Khaleghabad	Shahid Bahonar University of Kerman
Nahid Jafari	Shahid Bahonar University of Kerman
Abazar keikha	Velayat University
Farshid Keynia	Kerman Graduate University of Technology
Amid Khatibi	Islamic Azad University
Mohsen Khosravi	Shahid Bahonar University of Kerman
Marjan Kuchaki Rafsanjani	Shahid Bahonar University of Kerman
Fatemeh Kouchakinejad	Iranian Journal of Fuzzy Systems
Alimohammad Latif	University of Yazd
Malihe Maghfoori Farsangi	Shahid Bahonar University of Kerman
Najme Mansouri	Shahid Bahonar University of Kerman
Mashaallah Mashinchi	Shahid Bahonar University of Kerman
Mina Mirhosseini	Higher Education Complex of Bam
Abousaleh Mirzaei	Higher Education Complex of Bam
Hassan Mishmasht Nehi	University of Sistan and Baluchestan
Mohammad Molaei Emamzadeh	Shahid Bahonar University of Kerman
Somayeh Motamed	Islamic Azad University
Ardavan Najafi	Islamic Azad University
Ali Naserasadi	Higher Education Complex of Zarand
Komail Nekooei	Shahid Bahonar University of Kerman
Akbar Paad	University of Bojnurd
Abbas Parchami	Shahid Bahonar University of Kerman
Zeinab Khatoun Portaheri	Higher Education Complex of Bam
Jafar RazmAra	University of Tabriz
Akbar Rezaee	Payam-noor University of Kerman
Milad Riyahi	Shahid Bahonar University of Kerman
Azim Rivaz	Shahid Bahonar University of Kerman
Reza Saadati	Iran University of Sci. and Technology
Amir Sabbagh Molahosseini	Islamic Azad University
Ali Shakiba	Vali-e-Asr University of Rafsanjan
Ayoub Sheikhi	Shahid Bahonar University of Kerman
Sadegh Sulaimany	University of Kurdistan
Mohadese Soleimanpour-moghadam	Higher Education Complex of Bam
Hamed Tabrizchi	University of Tabriz
Seyyed Mahmoud Taheri	University of Tehran
Mohammad Teshnehlab	K. N. Toosi University of Technology
Seyyed Ali Torabi	University of Tehran
Mohammad Ali Yaghoobi	Shahid Bahonar University of Kerman
Fahimeh Yazdanpanah	Vali-e-Asr University of Rafsanjan
Saeedeh Zahiri	Higher Education Center of Eghlid
Jafar Zanjani	Farhangian University of Zanjan

Keynote Speeches and Workshops

CFIS2022

Keynote Speech Title	Time
Evolutionary Intelligence in (Big) Data Analytics and Optimization **Prof. Amir H. Gandomi, University of Technology Sydney, Australia**	Wednesday March/2/2022 10:30-11:30

Abstract
Evolutionary Computation (EC) has been widely used during the last two decades and has remained a highly-researched topic, especially for complex engineering problems. The EC techniques are a subset of artificial intelligence, but they are slightly different from the classical methods in the sense that the intelligence of EC comes from biological systems or nature in general. The efficiency of EC is due to their significant ability to imitate the best features of nature which have evolved by natural selection over millions of years. The central theme of this presentation is about EC techniques and their application to complex smart cities and infrastructures problems. On this basis, first I will talk about an evolutionary approach called genetic programming for data mining. Applied evolutionary computing will be presented, and then their new advances will be mentioned such as big data mining. Here, some of my studies on big data mining and modelling using EC and genetic programming, in particular, will be presented. Case studies' topics include forward design, and inverse design. In the second section, the evolutionary optimization algorithms and their key applications in the design optimization of complex and nonlinear engineering systems will be discussed. It will also be explained how such algorithms have been adopted to engineering problems and how their advantages over the classical optimization problems are used in action. Optimization results of large-scale engineering systems and many-objective problems will be presented which show the applicability of EC. Some heuristics will be explained which are adaptable with EC and they can significantly improve the optimization results.

Amir H. Gandomi is a Professor of Data Science and an ARC DECRA Fellow at the Faculty of Engineering & Information Technology, University of Technology Sydney. Prior to joining UTS, Prof. Gandomi was an Assistant Professor at Stevens Institute of Technology, and a distinguished research fellow in BEACON center, Michigan State University. Prof. Gandomi has published over three hundred journal papers and 12 books which collectively have been cited more than 26,000 times (H-index = 74). He has been named as one of the most influential scientific minds and Highly Cited Researcher (top 1% publications and 0.1% researchers) for five consecutive years, 2017 to 2021. He also ranked 17th in GP bibliography among more than 12,000 researchers. He has served as associate editor, editor, and guest editor in several prestigious journals such as AE of IEEE TBD and IEEE IoTJ. His research interests are global optimization and (big) data analytics using machine learning and evolutionary computations in particular.

Keynote Speech Title	Time
Overview of new families of fuzzy implications and their applications in fuzzy systems **Prof. Michal Baczynski, University of Silesia in Katowice, Katowice, Poland**	Wednesday March/2/2022 16:15-17:15

Abstract
The mathematical base of fuzzy systems is a multivalued logic, where among others, we use different generalizations of classical logic connectives. In our talk, we concentrate on analyzing multivalued implications called fuzzy implications in the literature. This class of functions plays a significant role both in theory and applications, as can be seen from their use in, among others, approximate reasoning, fuzzy control, image processing, data analysis and multivalued mathematical logic.
We will review the different classical families of fuzzy implications, focusing primarily on their main properties and possible characterizations. Next, we analyze some new families of fuzzy implication functions; in particular, we present the preference implications and new weighted and mean-implication functions. We will show why these implications can be helpful for engineers and specialists who design and implement fuzzy systems.

Prof. Michał Baczyński

Deputy Dean of the Doctoral School at the University of Silesia in Katowice, Vice President of the Upper Silesian Branch of the Polish Mathematical Society, Faculty of Science and Technology, University of Silesia in Katowice, Bankowa, Katowice, Poland

Prof. Michał Baczyński deals with the mathematical foundations of intelligent systems, especially fuzzy systems. He analyses different methods and systems where multi-valued operators are used. His main object of scientific interest are fuzzy (multi -valued) connectives, in particular fuzzy implication functions, used in many different fields of computer science and computational intelligence.

Keynote Speech Title	Time
The Sound of Health **Prof. Mohammad-R Akbarzadeh-T, Ferdowsi University of Mashhad (FUM), Iran**	Wednesday March/2/2022 19:30-20:30

Abstract

The recent Covid19 pandemic gave modern technologies an opportunity to showcase their relevance in serving humanity. The unprecedented availability of inexpensive processing power, wide-scaled network connectivity, and the extraordinary variety of sensing devices enabled many to work remotely, study in virtual classrooms, and buy their necessary everyday goods without leaving home. Nevertheless, the pandemic caused many patients to avoid hospitals and medical facilities when they needed it the most. It also contaminated many patients who did seek help but were healthy otherwise. Telemedicine, a relatively well-established concept long before the pandemic, shows great potential in this regard.

In this talk, I will present the utility of sound, particularly the human voice, as the means for inexpensive and readily available telemedicine technology for remote medical screening. Besides a great utility during pandemics, this means that it could offer increased accessibility to a wide range of underprivileged patients in far and remote areas. However, realistic constraints of low-quality microphones and the narrow bandwidth of transmission lines present significant challenges in this process. The available features are also many, their signal is noisy, and there is considerable uncertainty in mapping these features with the disease. This leads us to the role of intelligent systems for automatic screening, such as fuzzy logic for feature selection and deep neural network structures for signal classification. In particular, I hope to review our recent work in this domain using intelligent feature selection, classification, and testing/training strategy in diagnosing Diseases such as Parkinson's, Covid29, and several other lung diseases.

Mohammad-R. Akbarzadeh-T. (Ph.D. University of New Mexico (UNM), 1998) is currently a professor and director of the Center of Excellence on Soft Computing and Intelligent Information Processing (SCIIP), Electrical Engineering Department, Ferdowsi University of Mashhad. From 1996-2003, he was also with the NASA Center for Autonomous Control Engineering at UNM. In 2006 and 2017, he was also with the Berkeley Initiative on Soft Computing (BISC), UC Berkeley as a visiting scholar. In 2007, he also served as a consulting faculty at the Department of Aerospace and Aeronautic Engineering, Purdue University.

Prof. Akbarzadeh is the founding president of the Intelligent Systems Scientific Society of Iran, the founding councilor representing the Iranian Coalition on Soft Computing in IFSA, an IEEE senior member, and the founding faculty councilor of the IEEE student branch until 2008. He has received several awards, including the Caro Lucas Life Achievement Award from the IEEE Iran Section in 2021, the Outstanding Faculty Award from Ferdowsi Academic Foundation in 2021, the National Outstanding Faculty Award in 2019, and the IDB Excellent Leadership Award in 2010. He has also won the FUM university's Outstanding Faculty Award three times. His research interests are bio-inspired computing/optimization, fuzzy logic and control, soft computing, multi-agent systems, complex systems, robotics, cognitive sciences, and medical informatics. He has published over 450 peer-reviewed articles in these and related research fields.

Keynote Speech Title	Time
Role of Intelligent systems in the Interpretation of Deep Neural Networks in a Post-pandemic World **Prof. Saman K. Halgamuge, University of Melbourne, Australia**	Thursday March/3/2022 10:30-11:30

Abstract

Popular models of AI, in particular machine learning based models have three significant deficiencies: they are mostly manually designed using the experience of AI-experts; they lack human interpretability, i.e., users do not understand the AI architectures either semantically/linguistically or mathematically; and they are unable to dynamically change when new data are acquired. Addressing these deficiencies would provide answers to some of the valid questions about traceability, accountability and the ability to integrate existing knowledge (scientific or linguistically articulated human experience) into the AI model. This keynote addresses these deficiencies in the context of major global problems in the post pandemic world and how intelligent systems can contribute to finding solutions.

Prof. Saman Halgamuge received the B.Sc. Engineering degree in Electronics and Telecommunication from the University of Moratuwa, Sri Lanka, and the Dipl.-Ing and Ph.D. degrees in data engineering from the Technical University of Darmstadt, Germany. He is currently a Professor of the Department of Mechanical Engineering of the School of Electrical Mechanical and Infrastructure Engineering, The University of Melbourne, Australia. He is also an honorary professor of Australian National University. He is a Fellow of IEEE, a distinguished Lecturer of IEEE Computational Intelligence Society and listed as a top 2% most cited researcher for AI and Image Processing in the Stanford database. His research interests are in AI, machine learning including deep learning, optimization, big data analytics and their applications in biomedicine and engineering.

Keynote Speech Title	Time
Data-Driven Fuzzy Modeling **Prof. Irina Perfilieva, University of Ostrava, Czech Republic**	Thursday March/3/2022 16:15-17:15

Abstract
The talk will focus on efficient data-driven modeling associated with the inverse problem and feature extraction. We show how the theories of manifolds and F-transforms contribute to these delineated areas. The manifold hypothesis states that the shape of observed data is relatively simple and that it lies on a low-dimensional manifold embedded in a higher-dimensional space.
We contribute to the problem of manifold learning. We show that a space whose topological structure is characterized by a fuzzy partition naturally leads to so called Riemannian spaces or Riemannian manifolds. Finally, we show how the discussed notions contribute to the mathematics of deep learning.

Prof. Irina Perfilieva, PhD, dr.h.c., prof.h.c., Centre of Excellence IT4Innovations division of the University of Ostrava, Institute for Research and Applications of Fuzzy Modeling, Ostrava, Czech Republic

Research Activities:

- fuzzy approximation and fuzzy topological spaces;
- fuzzy transforms, image processing, mathematical morphology;
- fuzzy differential equations;
- fuzzy logic, approximate reasoning;
- systems of relation equations and their solvability.

For her long-term scientific achievements, she was awarded on the International FLINS 2010. She received the memorial Da Ruan award in 2012. Since 2013, she is an EUSFLAT Honorary Member, and since 2019 she is an IFSA Fellow.

Keynote Speech Title	Time
Some recent extensions of fuzzy integrals applied to the computational brain problem **Prof. Humberto Bustince, Universidad Publica de Navarra, Spain**	Thursday March/3/2022 19:30-20:30

Abstract
We start revising some extensions of the classical Choquet and Sugeno integrals that have appeared in the literature and which replace sums, products, min or max by more general operators in their definition. The resulting functions are not, in general, aggregation functions, but they belong to the wider class of pre-aggregation functions, for which monotonicity is required only for some (and not every) positive direction. We discuss the utility of these functions in some ensemble techniques in order to fuse data from different channels in order to be able to predict from electroencefalographical signals whether a given subject is thinking on the movement olf one hand or another. Experimental result with these new families of functions improves those obtained when the classical Choquet and Sugeno integrals are obtained.

Humberto Bustince Sola is a full professor of Computer Science and Artificial Intelligence at the Public University of Navarra and honorary professor at the University of Nottingham since 2017. He is the main researcher of the Research Group on Artificial Intelligence and Approximate Reasoning, whose research lines are both theoretical (data fusion functions, information and comparison measures, fuzzy sets and their extensions) and applied (Deep learning, image processing, classification, machine learning, data mining, big data or the computational brain). He has led 13 research projects funded by national and regional governments, and two excellence networks on soft computing.
He has been the main researcher in projects with companies and entities such as Caja de Ahorros de Navarra, INCITA, Gamesa Tracasa or the Servicio Navarro de Salud. He has taken part in two international projects. He has authored or coauthored more than 300 works, according to Web of Science, including around 160 in Q1 journals. He was a highly cited researcher among the top 1%most relevant scientists in the world in 2018, according to Clarivate Analytics. He collaborates with first line research groups from countries such as United Kingdom, Belgium, Australia, the Czech Republic, Slovakia, Canada or Brasil. He is editor in chief of the Mathware&Soft Computing online magazine of the European Society of Fuzzy Logic and technologies and of the Axioms journal. Associated editor of the IEEE Transactions on Fuzzy Systems journal and member of the editorial boards of the journals Fuzzy Sets and Systems, Information Fusion, International Journal of Computational Intelligence Systems and Journal of Intelligent & Fuzzy Systems. Moreover, he is a coauthor of a book about averaging functions, and has been the co-editor of several books. He has been in charge of organizing several first level international conferences such as EUROFUSE 2009 and AGOP 2013. He is Senior Member of IEEE y Fellow of the International Fuzzy Systems Association (IFSA). Member of the Basque Academy of Sciences, Arts and Literature, Jakiunde, since 2018. He has advised 11 Ph.D thesis.

He was awarded the Cross of Carlos III the Noble by The Government of Navarra in 2017. He got the National Computer Science Prize José García Santesmases in 2019 and the Scientific Excellence Award of EUSFLAT the same year.

Keynote Speech Title	Time
Possibility-Based Perception: A Reflection on Recognition of Shape Pattern **Prof. Kambiz Badie, University of Tehran & ITRC, Iran**	Friday March/4/2022 10:30-11:30

Abstract
Both "probability" and "possibility" have long been considered as key items for managing uncertainty in problem-solving / decision-making issues like classifying different types of patterns. While the major concern of probability is to emphasize on the observations from experiencing different cases, possibility mostly emphasizes on the aspect of structurality with regard to a situation. In this way, the necessary condition for justifying the utility of possibility for managing uncertainty within the process of pattern classification is the ability of considering structurality for the ongoing patterns. It is to be noted that, the structurality of a pattern or a situation goes back to the fact that how its parts or components can comply with certain models with certain characteristics. In this way, "possibility-based recognition" is concerned with a range of possibilities with regard to a variety of significant substructures in a pattern, which come together to give sense/identity to that pattern or situation on the base of the models belonging to these substructures. It is believed that such an approach to recognition takes place on the base of a kind of information fusion whose function is to incorporate a number of possibilities regarding the existing substructures in a pattern in order to minimize plausibly the very uncertainty which exists within the process of pattern classification. The role of fusion in our approach is that, once the vector representing a pattern was constructed on the base of the information on average possibility degrees for a variety of significant substructures in that pattern, it is first tried to determine the value of similarity between the components of this vector and the corresponding significant substructures in the existing classes taking into account a kind of similarity function and then determine the products of these similarity values for each class. Obviously, the class for which the value of this product is the most, should be considered as the most suitable class for the corresponding pattern. Fusion of possibility degrees can be performed at two levels; (i) characterizing the significant substructures in a pattern with the purpose of determining the most reasonable class for it, and (ii) final classification of a pattern including multiple classes.

As examples of patterns with structural nature, we may mention "patterns of shapes" consisting of a variety of curve segments or curve/region boundaries, "patterns of semantic entities" containing lexicons with certain meanings, and "patterns of mathematical/logical expressions" containing logical symbols, predicates, arguments, functions and variables. What we try to clarify in this presentation is how "possibility-based recognition" in the way explained above has the ability to perform a plausible classification of "patterns of shapes" taking into account the role of information fusion. Within the scope of shapes, some examples are discussed regarding classification of both Latin and Arabic characters, mentioning how based on the possibility degrees belonging to the corresponding substructures in these characters, their classification can be performed in a plausible manner. Shapes in general include "objects" (living, non-living and artificial), phenomena (cosmic, social, organizational and personal) and symbols (alpha-numeric and conventional configurations including message).

As a conclusion, it is significant to notice that the possibilistic approach to recognizing structurally-rich patterns provide a conducive opportunity to consider only those significant substructures, which are peculiar to the identity of the classes, thus avoiding the information regarding the irrelevant substructures that makes the process of classification rather time-consuming. Interestingly, the process of intuition in human being, which is a sort of high-level perception, seems to approve this fact.

Key Words—Perception, possibility, uncertainty, pattern classification, fusion, substructures, shape.

Keynote Speech Title	Time
Derivative of fuzzy functions and its application in fuzzy differential equations **Dr. Alireza Khastan, Department of Mathematics, IASBS, Iran**	Friday March/4/2022 15:15-16:15

Abstract

One of the important concepts in fuzzy mathematics is the concept of fuzzy derivative. In last 30 years, several definitions for fuzzy derivative have been presented.

In this talk, first we give a brief summary on different fuzzy derivatives including Hukuhara derivative, Generalized Hukuhara derivative and Metric derivative and we study some of their important properties. Next, we present the application of fuzzy derivative in fuzzy differential equations and some of the solution methods.

Workshops (In Persian)
9th Joint Congress on Fuzzy and Intelligent Systems

Monday, Feb/28/2022
Fuzzy and Grey Mathematical Programming **Dr. Seyyed Hadi Nasseri and Davood Darvishi, University of Mazandaran** **Chaiman: Prof. Behrouz Fathi Vajargah**
Knowledge Graphs and their Applications in Intelligent Systems **Dr. Amin Anjomshoaa, National University of Ireland Galway** **Chaiman: Dr. Marjan Kuchaki Rafsanjani**
Thursday, March/01/2022
How can you have a successful Monte Carlo simulation in calculations? **Prof. Behrouz Fathi-Vajargah, University of Guilan** **Chairman: Dr. Seyyed Hadi Nasseri**
Blockchain technologies, cryptocurrencies and smart contracts **Prof. Hossein Nezamabadi-pour, Shahid Bahonar University of Kerman** **Chairman: Dr. Mahdi Eftekhari**
Monday, March/14/2022
Fuzzy Sets and Fuzzy Logic **Professor Seyyed Mahmoud Taheri, University of Tehran** **Chaiman: Dr. Me'raj Abdi**

List of English Papers

#Page	Paper Title
	Presented Papers
38	A Dynamic SDN-based Privacy-Preserving Approach for Smart City Using Trust Technique
43	A Fault Isolation Approach for Data-Driven Device Replacement Decision Making
49	A Fractional Multi-Commodity Network Flow Problem with Uncertain Multi-Choice Coefficients: Model Formulations and Solution Approach
55	A New Approach for Solving the Interval Type-2 Fuzzy Transportation Problem Based on interval Linear Programming Problems
61	A new contrast enhancement method for Color dark and low-light images
68	A New Edge and Pixel-Based Image Quality Assessment Metric for Colour and Depth Images
78	A Novel Approach to Optimizing the Initial Path of Mobile Robots in Static Environments
85	A Novel Control Strategy Based on Fuzzy Logic in Islanded Microgrid
92	A novel ensemble feature selection method through Type I fuzzy
98	A Secure mechanism to protect UAV communications
104	Adaptive TSK Fuzzy Terminal Sliding-Mode Control of Two Coupled Cart-Mounted Inverted Pendulums
110	Adjoint relations of S-fuzzy posets with some categories
114	An Aggregated Revenue-Driven VPP Model Based on Marginal Price Tracking for Profit Maximization
121	An Efficient Collaborative Filtering for Recommendation Systems Using Differential Machine Learning
127	An efficient outlier detection method based on distance ratio of k-nearest neighbors
132	An Ensemble of Feature Selection Algorithms Using OWA Operator
138	An Ensembled Text-based Multi-Stacked-CNN-BILSTM for Rumor Detection on Social Media
144	Application of fuzzy generalized power series for the fuzzy linear inhomogeneous differential equation

263	How weight-sharing mechanisms affect the performance of deep Siamese networks
269	Hyper Ideals in Hyper Equality Algebras
274	Improved LOF Algorithm Using Random Point
280	Improving Image Captioning with Local Attention Mechanism
286	Improving the Obtained Results of Monte Carlo Simulation for Interval Linear Programming Problem by Using Particle Swarm Optimization
290	Intelligent fault detection of planetary gearbox using vibration signal processing by empirical mode decomposition and an integrated artificial neural network-support vector machine classifier
296	Intelligent Vibration-based Anomaly Detection for Electric Motor Condition Monitoring
301	Intuitionistic Fuzzy Multiset finite Subautomata
305	Jaccard Pseudo-Similarity of Fuzzy Parameterized Fuzzy Soft Matrices and Its Application to Diagnosis of Parkinson's Disease
310	Lattice-Valued fuzzy hyper ideals in hyper residuated lattices
314	Latticized-Hamacher optimization problem subject to fuzzy relational equations
318	Learning-based fuzzy c-means clustering using mixtures of Student's-t distributions with missing information
324	Location of fire station in Bam city using Fuzzy Analytic Hierarchy Process
331	Minimizing a composite fuzzy function in terms of subgradient
336	Modified Relay Node Placement in dense 3D Underwater Acoustic Sensor Networks
341	Multiobjective Zero-Sum Games in Intuitionistic fuzzy Environment
345	New results on prime A-ideals in MV –semimodules
349	Notes on energy of matching in fuzzy graphs
353	Numerical solution for Interval Initial Value Problems based on interactive arithmetic
360	On P-torsion EQ-modules and P-cyclic EQ-modules
364	On the Completeness of a quantale-valued metric space
368	Online streaming feature selection based on Sugeno fuzzy integral
374	Power Allocation in CRNs based on QoS and QoE
381	Reinforcement Learning Reward Function for Test Case Prioritization in Continuous Integration
387	Residuated lattice congruences via directed kernels

391	Scattering and Regional Features Fusion Using Collaborative Representation for PolSAR Image Classification
397	Soaker ideals in MV –algebras
401	Solving fuzzy bilevel linear programming problem based on interval approximation
405	Some Properties of L-graphs
411	Some Results in Projective System of BL-algebras
415	Special types of NM-algebras
420	The Analysis of a Fractional Network-Based Epidemic Model with Saturated Treatment Function and Fuzzy Transmission
430	The Behavior of L-graph Automata
435	The characterization of $\boldsymbol{L^B}$ −valued GFA via $\boldsymbol{L^B}$ −valued operators
441	The Decomposition Theorems for Residuated Lattices via Directed Kernels
446	The fuzzy D'Alembert solutions of the fuzzy wave equation under generalized differentiability
455	The new modifications of distance measures on hesitant fuzzy numbers
459	True-False triangular norm and conorms and its Application
465	Weak solutions to fuzzy stochastic differential equations under sub-fractional Brownian motion
469	Zero sets in MV −algebras of continuous functions
473	Asymmetric distributions based on the t-copula
476	Improving the fit for diagonal copula based on Kendall's tau and tail dependence preserving transformation
481	Multivariate ageing intensity
482	Testing positive quadrant dependence with discrete copulas
483	On quantifying and estimating directed dependence
484	Invariant copulas under univariate truncation
Poster Papers	
486	A Nitrate Enzymatic Biosensor based on Optimized Machine Learning Techniques
491	ECG-Based Prediction of Epileptic Seizures Using Machine Learning Methods
498	Fractional entropy and its applications in fuzzy c-means clustering
502	Fuzzy stability of involutions via fixed point technique
507	Intelligent detection of bone fractures using data mining and image processing methods

Presented Papers

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)
March 2-4, 2022, Higher Education Complex of Bam

A Dynamic SDN-based Privacy-Preserving Approach for Smart City Using Trust Technique

Jafar A. Alzubi[1], AliAkbar Movassagh[2], Mehdi Gheisari[3], Hamid Esmaeili Najafabadi[4], Aaqif Afzaal Abbasi[5], Yang Liu[6], Zhou Pingmei[7], Mahdieh Izadpanahkakhk[8], AmirHossein Pourishaban Najafabadi[9]

[1] *Faculty of Engineering, Al-Balqa Applied University, Salt - Jordan*
[2]*Department of Medical Physics and Biomedical Engineering, School of Medicine, Tehran University of medical sciences, Tehran, Iran*
[3]*Young Researchers and Elite Club, Parand Branch, Islamic Azad University, Parand, Iran*, Mehdi.gheisari61@gmail.com
[4]*University of Calgary, Canada*
[5] *Department of Software Engineering, Foundation University, Islamabad 44000, Pakistan*
[6]*Department of computer science and technology, Harbin Institute of Technology, Shenzhen, China*
[7]*Shenzhen BKD Technology Co., Ltd, Shenzhen, China*
[8]*DIET department, La Sapienza University, Rome, Italy*, izadpanahkakhk@diag.uniroma1.it
[9]*Department of Electrical and Computer Engineering, University of Yazd, Yazd, Iran*

***Abstract*—A smart city is an Internet-based application of things that automates city management with no need for human interference. Exchanging data via devices obviate some challenges in intelligent cities. In a smart city, Internet-of-Things (IoT) devices may detect sensitive data, posing a risk of privacy violation and system harm. We discover that existing solutions are either too expensive or ineffective at limiting unintended disclosure of sensitive data to build a dependable, smart city. The fact that they create static surroundings is the fundamental reason behind this. Software-Defined Networking (SDN) technology has recently evolved to configure the network for performance and monitoring improvement. This study offers a work-in-progress that uses the SDN to protect the privacy of IoT devices by creating a dynamic SDN-based privacy-preserving ecology. The mechanism of the SDN controller performs under the nodes' mutual trust; it chooses various routes from the IoT device to the Cloud space destination dependent on the level of confidence. The packet is re-routed if the SDN controller identifies a device that does not trust its neighbor. Then it instructs the owner to deliver data over a different path. To demonstrate its improved performance, we are currently evaluating it from the perspective of overhead criteria in the future.**

***Index Terms* —Internet of Things; Trust, Smart City**

I. INTRODUCTION

Due to the growing number of inhabitants worldwide by 2050, city management is one of the motivational issues. As predicted, fresh practical solutions to efficiently run cities in duties such as public transit, waste collection, and so on are in high demand [1]. The Internet of Things (IoT) and cloud computing are two developing technologies that have the potential to improve existing city management. A city can address its difficulties more intelligently by combining these two technologies, which is often referred to as a Smart City.

The Internet of Items envisions a connected universe of existing things that can be accessed at any time and from any location. Cloud computing, on the other hand, refers to all of the computer gear and software that a third-party corporation provides to the end-user as a service over the Internet.

Because of the widespread use of IoT, the number of IoT devices integrated into an intelligent city will number in the billions by 2050 according to the Gartner's report [2]. The intelligent city market will be worth USD 1.2 trillion by the end of 2020, which is nearly 3 times greater than in 2014[3]. Within a smart city, IoT sensors drive data to cloud computing not only to provide more in-depth analysis but also to share intelligence

*Corresponding Authors: Mehdi Gheisari, Yang Liu

alongside other people [4]. An example of an IoT application is a smart house, whereas a traditional home is outfitted with IoT with the goal of improved management. Fig. 1 displays a smart home, which is one of the key components of a smart city.

Figure 1. Example of an old-fashioned intelligent home in a modern city.

As illustrated in Fig. 1, IoT technology in an intelligent city (like intelligent houses), share data with the cloud computing space to make top-quality decision-making [5]. These devices can be mobile, such as cars, or stationary, such as smartwatches, which stream information straight to the cloud.

More data, in general, means more opportunities to provide high-level services. However, it also brings up a significant challenging issue, which is privacy protection since that some smart city device may produce sensitive data. Malicious behaviors may create trouble and errors in the system if sensitive data is released involuntarily. As a result, while they collaborate to provide high-level services, these data should be securely protected. If the privacy concern is addressed, citizens will be more inclined to use the Internet of Things-based smart city. There are three sorts of sensitive data that should not be accidentally published:

1) Personal information: information that allows an individual or an IoT device to be directly authenticated, such as a person's name, the number of people alive in a large building, a low brake oil level in an autonomous vehicle, a social security number, and so on.

2) Semi-sensitive information, such as wages and medical conditions.

3) Information with a quasi-identifier: We must avoid unintentional publication of this type of data because by integrating and merging sensed data with information collected from external sources, we may distinguish the data of a given IoT device or individual. such as a public-voting registration database, a database of IoT device locations, a database of hospital registrations, and so on. In other words, a specific IoT device can be identified by combining the provided data with external databases.

Considering the above, we should provide methods to avoid sensitive data from being accidentally disclosed. Otherwise, there is a good chance it will be abused by adversaries attempting to physically or cyber-attack the system. Assume an IoT device detects the number of individuals alive in a significant building. This information must not be accidentally released since adversaries can alert a third party, who can then analyze the information and launch physical or virtual attacks on the building.

We discovered that present solutions could not provide an effective environment because they are either too expensive [6], [4], or lack the required performance from a privacy-preserving standpoint [7]. Finding a middle ground between these two research objectives is a big difficulty in creating an efficient smart city where inhabitants can trust smart city technologies while knowing that sensitive data will not be accidentally released.

Meanwhile, the Software-Defined Networking approach has become a popular networking model for dividing the data plane and the control plane [8]. This split results in a network that is flexible to administer. It provides a centralized view of a full network and makes network management more efficient. In other words, it enables network equipment administration from a single centralized controller, which is impossible to achieve with the classic Simple Network Management Protocol (SNMP). In an SDN environment, one of the key advantages is the ability to control data traffic. Network setup and maintenance will be done in a more convenient manner using SDN [9-15].

In this paper, we offer a privacy-preserving approach for IoT devices integrated into a smart city to address some challenges between current studies., where the packets are through the SDN controller, data is delivered to the cloud computing area. They transmit their data to the cloud computing space for easier data exchange and high-level decision-making since it delivers on-demand computer system resources [10,16-26]. If neighbors do not trust each other, this is accomplished by rerouting data packets.

We intend to analyze the proposed method and compare it to the state-of-the-art after publishing this work-in-progress. We will evaluate it based on many parameters, including the quantity of privacy-preserving degree, computing cost, latency, and communication overhead. In a nutshell, the proposed technique entails the stages below:

1) To operate the network flexibly, we equip contemporary smart cities with the SDN paradigm.

2) We present a solution for preserving the privacy of IoT devices on top of the outfitted environment. If the IoT device's trust in its next neighbor is less than 50%, the SDN controller instructs it to reroute its sensitive data. This schema is based on a guess. To locate the best deal, more research is required. As a result, if the device generates sensitive data, the SDN controller is in charge of routing the data from the IoT device to itself. The controller does not allow data to be redirected from the established path if a node's trust in its neighbor is less than 50%. Rather, it instructs the IoT device to send data over a new route defined by the SDN controller. On the other hand, if the trust level exceeds 50%, sensitive data will be sent via a predetermined channel. Finally, the SDN controller transfers the data to Cloud Computing for additional analysis.

3)We will assess our proposed method from multiple perspectives, including privacy-preserving degree and penetration rate, overhead, and latency, to demonstrate its improved performance over current studies. We will demonstrate that, although adding more overhead to the IoT-

based smart city, our technology effectively protects the privacy of IoT devices. If it can be widely employed in smart cities.

This is how the paper is structured. The second section discusses the literature that is relevant to this topic. The proposed approach is then thoroughly explained in Section III. Section V brings the paper to a close and makes recommendations for further research.

II. LITERATURE REVIEW

Several current research trends aim to use the SDN paradigm to protect the privacy of IoT devices in a smart city. The general concept is that data is sent to the cloud computing space for additional analysis and command execution. In [11], the authors suggested a paradigm for providing end-to-end security and privacy in 5G-enabled vehicle networks. Their suggested system used the SDN paradigm to simplify network management while achieving optimal network connectivity. It consists of two modules: the first is an authentication protocol that uses elliptic curve cryptography (ECC) for mutual authentication between cluster heads (CH) and certificate authorities (CA) in automotive contexts using the SDN [12]. The Intrusion Detection module is the second designed module, and it detects potential intrusions in the system. The module has minimal computational complexity, according to the researchers. They used three simulators to fully exploit the suggested framework's potential benefits (e.g., NS3, SUMO, and SPAN). The first module was evaluated based on its security features. The detection rate, false-positive rate, accuracy, detection time, and communication overhead were all included in the second module evaluation compared with the state of the art.

Using the SDN paradigm, the authors of [5] devised a solution for safeguarding the privacy of IoT devices in a smart city. Their technology is context-aware, allowing users to react to their surroundings based on their current situation. First, they installed the SDN paradigm in the smart city. Then they implement a privacy-preserving mechanism in which the device generates sensitive data and separates it into two portions, each comprising 70% and 30% of the original data. The first division is then sent to the SDN controller through the most secure path. The remaining data is sent to the SDN controller via a built virtual private network (VPN). The controller then aggregates the device's data. Finally, it sends the compiled data to the smart city cloud for further analysis and command execution. Several evaluation measures, including accuracy, penetration time, and overload, were used to assess their suggested technique. They also contrasted their answer to what was currently available. They discovered that their method is more effective at preventing unintentional sensitive data disclosure. Simultaneously, it adds to the smart city's workload.

Meantime, [13] inspires us, even though the authors did not use the SDN paradigm. They presented a lightweight privacy-preserving data aggregation (PDA) approach for the fog-computing-enabled IoT environment. To encrypt the data flow, they used the homomorphic Paillier cryptosystem. They also used the Chinese remainder theorem (CRT) to combine data from a variety of IoT devices [14]. To provide a more efficient solution, they also used a one-way hash chain function to filter injected bogus data at the network edge-level forging and had more efficient authentication of IoT devices [15]. They also used differential privacy-preserving (DP) as a supplement to create more effective privacy-preserving [16]. Furthermore, their technology is light enough to be used in real-time demanding situations. Aside from the benefits indicated, the PDA has a disadvantage in that it is not adaptable or agile.

Similarly, the authors of [2] devised a privacy-preserving approach for the IoT using the SDN paradigm. They installed software on the SDN controller, allowing it to govern IoT device data flow. To begin, the SDN controller divides IoT devices into numerous categories. Based on their associated class label, it then selects whether to encrypt, aggregate, or transfer their data to the SDN controller through an established VPN. Although the authors evaluated their proposed strategy based on the amount of overhead, this is insufficient to determine if their solution can effectively protect the privacy of IoT devices. Furthermore, the approach they provide is not context-aware.

In [8], the authors built on their earlier work and suggested a way for increasing the privacy of IoT-based smart cities with SDNs. They imagined a situation in which five smart buildings generate sensitive data and want to send it to a cloud computing environment for further study [17]. A solution was presented to maintain the privacy of smart buildings by splitting them into two sub-categories on top of the equipped smart city. An encryption mechanism is used if the smart building is classed in the first category. Otherwise, the data is split into two portions and sent to the SDN controller over two distinct routes.

They evaluated the amount of overhead and compared it to the time it takes IoT devices to deliver data straight to the cloud computing area to evaluate the proposed method. Although this strategy is cost-effective for many IoT devices, the solution's performance has not been evaluated in terms of privacy protection. Furthermore, encryption is used as a preventative step to conceal crucial data that can be omitted [18].

III. PROPOSED METHOD

A. Assumptions

We consider IoT devices already know whether the sensitivity of gathered data. Furthermore, specialists have already entered safe routes from an IoT device to the SDN controller into a database. As previously stated, the SDN controller knows the trust relationship between nodes via a database that is considered static in this paper.

B. Procedure

This section proposes a unique approach for improving the efficiency and privacy of IoT devices in a smart city. It includes the principle of dynamically preserving the privacy of IoT devices based on the degree of trust among nodes. Our scenario utilizes the same database as [5]. So far, we've got a laptop, a smartwatch, a smart building, a smartphone, a garbage can, and an autonomous vehicle. They are supplying numerical information. They intend to send their data to the Cloud. They have the SDN paradigm installed. As a result, IoT devices send data to an SDN controller first, then to the Cloud. As mentioned in the previous section, the SDN controller has the dataset of trust amounts between nodes. The SDN controller verifies the trust amount when an IoT device wants to send sensitive data. If the trust level is less than the threshold=0.5, the IoT device is

instructed to transfer sensitive data through a different random route. If not, it follows the predetermined path and transfers the data to its neighbor. The approach for the proposed solution in this paper is shown in the following.

Algorithm I. Proposed Privacy-preserving algorithm

```
Input: X= The sensed data
for all IoT devices do
    if X= Sensitive data then
      The SDN controller specifies the first secure
       route of the database from the IoT device to
       itself;
     for All Middle Nodes do
        T=The amount of trust among nodes;
        if T is less than 50% then
            SDN controller specifies a new random
             route;
Output: NULL
```

IV. ANALYSIS AND RESULTS

This section evaluates our proposed method from the overhead point of view. In other words, in this section, we determine how much strain it places on IoT devices or how much overhead the system should tolerate from the computational cost perspective.

Figure 2. The amount of overhead in terms of time

As Figure 2 shows, our solution imposes, in the beginning, around 35% overhead. This amount decreases during the steady phase.

CONCLUSION

In this ongoing research, we first installed the SDN paradigm in a smart home. Then, on top of the SDN controller, we installed software to manage data flow flexibly based on the amount of trust among nodes. If the level of trust is less than 50%, it instructs the IoT device to reroute its data and send it via a different secure route. However, if the percentage is greater than 50%, the device sends sensitive data via the predefined route. To complete this research, we intend to evaluate it from a variety of perspectives, including privacy preservation, overhead, latency, and so on. In the future, we want to expand this approach and examine it from many angles to see if it outperforms current studies.

ACKNOWLEDGMENT

This work is supported by the Shenzhen Basic Research (General Project)(No. JCYJ20190806142601687) and Shenzhen Stable Supporting Program (General Project) (No. GXWD20201230155427003-20200821160539001).

REFERENCES

[1] R. Du, P. Santi, M. Xiao, A. V. Vasilakos, and C. Fischione, "The sensible city: A survey on the deployment and management for smart city monitoring," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp.1533–1560, 2018.

[2] M. Gheisari, et al, "A method for privacy-preserving in IoT-SDN integration environment," in *2018 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Ubiquitous Computing & Communications, Big Data & Cloud Computing, Social Computing & Networking, Sustainable Computing & Communications* (ISPA/IUCC/BDCloud/SocialCom/SustainCom). IEEE, 2018, pp. 895–902.

[3] P. Neirotti, A. De Marco, A. C. Cagliano, G. Mangano, and F. Scorrano, "Current trends in smart city initiatives: Some stylised facts," *Cities*, vol. 38, pp. 25–36, 2014.

[4] M. Gheisari, et al, "Eca: an edge computing architecture for privacy-preserving in IoT-based smart city," *IEEE Access*, vol. 7, pp. 155 779–155 786, 2019.

[5] M. Gheisari, G. Wang, W. Z. Khan, and C. Fernández-Campusano, "A context-aware privacy-preserving method for IoT-based smart city using Software Defined Networking," *Computers & Security*, vol. 87, p. 101470, 2019.

[6] P. G. V. Naranjo, Z. Pooranian, M. Shojafar, M. Conti, and R. Buyya, "Focan: A fog-supported smart city network architecture for management of applications in the internet of everything environments," *Journal of parallel and distributed computing*, vol. 132, pp. 274–283, 2019.

[7] M. S. Rahman, I. Khalil, M. Atiquzzaman, and X. Yi, "Towards privacy preserving ai based composition framework in edge networks using fully homomorphic encryption," *Engineering Applications of Artificial Intelligence*, vol. 94, p. 103737, 2020.

[8] M. Gheisari, G. Wang, S. Chen, and H. Ghorbani, "Iot-sdnpp: A method for privacy-preserving in smart city with software defined networking," in *International Conference on Algorithms and Architectures for Parallel Processing*. Springer, 2018, pp. 303–312.

[9] W. Braun and M. Menth, "Software-defined networking using openflow: Protocols, applications and architectural design choices," *Future Internet*, vol. 6, no. 2, pp. 302–336, 2014.

[10] T.-S. Chou, "Security threats on cloud computing vulnerabilities," *International Journal of Computer Science & Information Technology*, vol. 5, no. 3, p. 79, 2013.

[11] S. Garg, K. Kaur, G. Kaddoum, S. H. Ahmed, and D. N. K. Jayakody, "Sdn-based secure and privacy-preserving scheme for vehicular networks: A 5g perspective," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 9, pp. 8421–8434, 2019.

[12] P. Szczechowiak, L. B. Oliveira, M. Scott, M. Collier, and R. Dahab, "Nanoecc: Testing the limits of elliptic curve cryptography in sensor networks," in *European conference on Wireless Sensor Networks*. Springer, 2008, pp.305–320.

[13] R. Lu, K. Heung, A. H. Lashkari, and A. A. Ghorbani, "A lightweight privacy-preserving data aggregation scheme for fog computing-enhanced iot," *IEEE Access*, vol. 5, pp. 3302–3312, 2017.

[14] M. Jafari, J. Wang, Y. Qin, M. Gheisari, A. S. Shahabi, and X. Tao, "Automatic text summarization using fuzzy inference," in *2016 22nd International Conference on Automation and Computing (ICAC).* IEEE, 2016, pp. 256–260.

[15] M. Gheisari, "The effectiveness of schema therapy integrated with neurological rehabilitation methods to improve executive functions in patients with chronic depression," *Health Science Journal*, vol. 10, no. 4, p. 1, 2016.

[16] E. GhadakSaz, M. R. Amini, P. Porkar, and M. Gheisari, "Design, implement and compare two proposed sensor data storages named semhd and ssw," From Editor in Chief, p. 78, 2012.

[17] Q. Liu, G. Wang, X. Liu, T. Peng, and J. Wu, "Achieving reliable and

secure services in cloud computing environments," *Computers & Electrical Engineering*, vol. 59, pp. 153–164, 2017.

[18] S. Yu, G. Wang, and W. Zhou, "Modeling malicious activities in cyber space," *IEEE network*, vol. 29, no. 6, pp. 83–87, 2015.

[19] Ashourian, Mohsen, et al. "An Improved Node Scheduling Scheme for Resilient Packet Ring Network." *Majlesi Journal of Electrical Engineering* 9.2 (2015): 43

[20] Fakhimi, Esmaeil, et al. "Design Two Sensor Data Storages." International Conference on Advanced Computer Theory and Engineering, 4th (ICACTE 2011). *ASME Press*, 2011.

[21] Mehdi Gheisari, et al, A Method for Privacy-preserving in IoT-SDN Integration Environment, *16th IEEE International Symposium on Parallel and Distributed Processing with Applications (ISPA 2018)*, 11-13 Dec. 2018, Melbourne, Australia

[22] M. Gheisari et al., "A Survey on Clustering Algorithms in Wireless Sensor Networks: Challenges, Research, and Trends," *2020 International Computer Symposium (ICS)*, Tainan, Taiwan, 2020, pp. 294-299

[23] Rezaeiye, Payam Porkar, et al. "Agent programming with object oriented (C++)." *Electrical, Computer and Communication Technologies (ICECCT)*, 2017 Second International Conference on. IEEE, 2017.

[24] Gheisari, Mehdi, et al. "MAPP: A Modular Arithmetic Algorithm for Privacy Preserving in IoT." *Ubiquitous Computing and Communications (ISPA/IUCC), 2017 IEEE International Symposium on Parallel and Distributed Processing with Applications and 2017 IEEE International Conference on. IEEE*, 2017.

[25] Movassagh, *et al.* Artificial neural networks training algorithm integrating invasive weed optimization with differential evolutionary model. *J Ambient Intell Human Comput* (2021).

[26] Natarajan, Y., et al.: An IoT and machine learning-based routing protocol for reconfigurable engineering application. *IET Commun*. 00, 1– 12 (2021).

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

A Fault Isolation Approach for Data-Driven Device Replacement Decision Making

Mohammad Mahdi Kalantari, Hossein Yektamoghadam and Amirhossein Nikoofard

Electrical Engineering Department, *K. N. Toosi University of Technology*, Tehran, Iran, mohammadmahdi.kalantari@email.kntu.ac.ir
Electrical Engineering Department, *K. N. Toosi University of Technology*, Tehran, Iran, Yektamoghadam@email.kntu.ac.ir
Electrical Engineering Department, *K. N. Toosi University of Technology*, Tehran, Iran, a.nikoofard@kntu.ac.ir

***Abstract*— Despite the ubiquity and the prominence of device replacement decisions, the process is not always as simple as it seems. Especially as data-driven approaches are becoming more vital and intelligent systems have recently outperformed the classical techniques. The objective of this paper is to propose a solution to provide the set of decisions for replacing a whole asset. To accomplish this, first, we translate the deterioration level into the monetary scale via a data-driven approach. This is done by modeling the aging effects via fault occurrences; then utilizing fault isolation techniques for the determination. Finally, the set of decisions can be made by the optimization procedure.**
In the end, to implement the proposed solution, an actual fault isolation data set is reproduced to form a pertinent data set for a fictitious firm. Then an ANN is exploited for the fault isolation, and a game-theoretic approach comprising GA is utilized for the decision-making part.

***Index Terms*—Device replacement, data-driven, ANN, GA, decision making, game theory.**

I. Introduction

A. Motivation

All devices are prone to degradation. This incidence is partially compensated by maintenance, but indeed it reaches a level for replacement. In many cases, making mistakes regarding equipment replacement can cost the company much more than all the cost savings achieved in all other areas of production and planning [1]. However, determining the optimal decisions to prevent extra costs is not always straightforward.

The recent phenomenal achievements of AI intrigue researchers to ponder its exploitation to facilitate their own complex problems. One of them can be this very field. AI gives us the capability to bring many new data into the picture. With the ability to interpret them, the situation will be changed.

Reaching a conclusion for the replacement decision making must be done by considering the conditions of the market, as well as the present and the candidate device(s). Thus a proper replacement suggestion is the conjunction of multiple decisions. These assessments must be thorough, verifiable, and preferably with an acceptable complexity.

In order to strike the right balance between these qualities, we are enticed to devise a viable AI-based approach.

B. Literature Review

Commencing by the broadest vision, selecting machinery is an inextricable subset of production planning, scheduling, and control.

From the mentioned vast field, by focusing on the improvement of the product quality as well as amending the production process, the discussion can be narrowed down to the optimization of the production process [2].

By reaching the traditional methods such as six sigma strategies to their limits in the optimization of the production process [3], machine learning approaches play a pivotal role in this area. As a few examples, Bayesian optimization in robotics [4], ANN in electroplating [5], GA in welding [6], multi-objective optimization algorithms in analog active filter design [7], and fault detection in thermoelectric energy harvesting devices [8] can be mentioned.

Since then, the replacement model has been used to evaluate a wide range of equipment, including forklifts, automobiles, medical ventilator equipment, buses, escalators, and aircraft [9, 10]. In this very field, operating and maintenance (O&M) costs are generally reliant on the age of the machine [10].

Among related methods papers, [11–13] propose a data reduction-based approach for classification of initial data and applying it to the case study devices.[14] describes an artificial neural network-based approach for the prediction of biochemical application. [15] utilizes ML for aging monitoring and lifetime prediction of IoT devices. [16] provides a novel algorithm by applying the Nash-Cournot model to reach the Nash equilibrium point. [17] defines a fault identification game on hybrid systems, and using two players: the environment and the identifier. [18] proposes an optimal pricing decision strategy using two nested Stackelberg games.

C. Contributions and Outline

Contributions of this paper can be separated into different parts.

- Measuring the deterioration level: This paper proposes the data-driven modeling of the aging effect via incidences detectable by fault isolation techniques. Defining these incidences might slightly deviate from the regular sense of *fault*, as mentioned in Section III, and it can comprise more aspects. In this way, the deterioration level of devices can be measured utilizing sensory data (which may not be observable usually) and become a basis for decision making.
- Translation to monetary scale: By expressing the aging effect in financial terms, we can define a unified cost function that encompasses all aspects of the problem: from devices to the market. Changing the work to suit other needs is therefore possible since money is a universal language and easily defined.
- Providing the best solution: In this work, finding the best *set of decisions* for replacing a device is elaborated. It cannot be called the *optimal* solution due to the inherent uncertainty of the problem. But it minimizes a cost function as in finding the optimal point. The mentioned set of decisions comprises three factors, as an example in this work. However, it can easily be extended in much the same way via updating the cost function.

The remainder of the paper is organized as follows. First, the problem is defined verbally and then formulated formally in Section II; besides, the main underlying assumptions are specified. In Section III, our solution is proposed, which is simulated in Section IV utilizing an actual data set to create a fictitious process, exploiting an ANN for fault isolation, and a game-theoretic approach as a decision-making tool. Finally, the work is concluded and future work is proposed in Section V.

II. Definition and Modeling the Problem

With a succinct description, a production firm is considered that incorporates a deteriorated device. Therefore, the questions would be *When should this device be revised?* and *Which device should it be replaced with?* to minimize the firm's cost. And if the production is limited to a specific period, as in our case, another question can be added as *How long should the production line be kept after the replacement?*

The solutions to these questions are provided under certain assumptions. The main ones are stated clearly, and the rest can be apprehended through the formulation.

A. Main Assumptions

- The demand function of the product is known and decaying at the ending times.
  A decreasing demand after reaching its peak is a realistic assumption. This embodies the time limit for the production that induces us to stop the production line after a while. And providing this demand function has its own techniques that it is assumed it is already done for the firm.
- Further deterioration is negligible in the decision time.
  It means the decaying of the demand function is faster than the deterioration of the device, and the state of the present and the next device are constant during the activity of the production period. Obviously, the device is needed to be replaced a maximum of one time during this period.
- The replacement consists of the entire asset.
  The provided solution regards the device as a whole and does not provide maintenance advice.

B. Formulation of the Problem

The problem of the device replacement can be regarded as an optimization problem. It is convenient to solve optimization problems through the definition of a cost or objective function. As it can be seen in [10], most approaches in this field are trying to minimize the cost rather than maximizing the profit. Thus a similar approach is chosen in this work. And for the dimension of the cost function, the monetary scale can be claimed as the most well-known language that approximately everything can be translated into.

The formal formulation can be commenced by simple subtraction, considering the inflation rate

$$J = expenditure|_{t_0} - revenue|_{t_0} \tag{1}$$

in which J indicates the cost, which consists of bringing the future transactions into the present time. Afterward, each element of (1) must be defined.

As a rudimentary model for the expenditure,

$$expenditure|_{t_0} = \left(\sum_{t=0}^{t_{replace}} C_F(d_1) \times \gamma^t \right) + P_{d_2} \times \gamma^{t_{replace}} + \left(\sum_{t=t_{replace}}^{t_{end}} C_F(d_2) \times \gamma^t \right) \tag{2}$$

can be defined in which C_F is the cost that faults of the device induce, that is a function of the current device (d_1), or the following device intended to be bought (d_2). And deterioration directly affects this part. The following notation, P_{d_2} denotes the price of the next device, and γ is the inflation rate, related to each period the prices are taken into account.

The next part of the cost function, revenue basically can be written as

$$revenue|_{t_0} = \left(\sum_{t=0}^{t_{replace}} P_P(d_1) \times D(t) \times \gamma^t \right) + P_{j_{d_1}} \times \gamma^{t_{replace}} + \left(\sum_{t=t_{replace}}^{t_{end}} P_P(d_2) \times D(t) \times \gamma^t \right) \tag{3}$$

that P_P indicates the price of the product produced by a specific device (deteriorated device might result in a product with lower qualities), and D is the demand function. The last notation, $P_{j_{d_1}}$, denotes the junk price of the present deteriorated device. If at the end of the production time, i.e.,

$t = t_{end}$, the replaced device is sold as second hand, another term can be added to (3) which is

$$P_{S_{d_2}}(t_{end}) \times \gamma^{t_{end}}. \tag{4}$$

In (4), $P_{S_{d_2}}$ denotes the second-hand price of the newly acquired device that is a decaying function of time.

As mentioned earlier, these are the rudimentary models that can become more sophisticated by adding terms relevant to the utilized case in a similar manner. But for preventing loss of generality, only a basic one is considered here.

III. Proposed Solution Methodology

The result of deterioration is the malfunctioning of the device. If enough sensory data is available, the malfunctioning can be detected.

By calling each type of malfunctioning a fault, we can utilize the fault isolation jargon to express and solve our problem. But it may deviate a little from the usual sense of the *fault*. For example, it can include occurrences such as damaging the product (which is not directly the fault of the machine), unscheduled stops, or breaking a part of the machine and demanding maintenance, or even overconsumption of the electricity.

Isolating these faults can be done via signal-based, model-based, or hybrid approaches. For further information, readers are referred to [19] and [20].

Thus after defining a set of faults that can be isolated (and adding more sensors if needed), the rudimentary approach will be assigning a price to each occurrence that forms the Cost Generation part in Fig. 1. This can also be more sophisticated if needed.

Up to here, we have the cost made by the deterioration of the device. Considering the profit of production and market parameters, it will be possible to make the replacement decision that minimizes the total cost or equivalently maximizes the total profit. This is done by searching the space created by parameters to find the extremums. This can be accomplished using PSO or GA, as two popular approaches.

But for having a proper result, this procedure is better to be done multiple times during a period. This means if the decision period is set to be one month, as an example, it is better to split recorded data into smaller parts, e.g., 2 days. In this given example, there would be 15 sets of data that the decision is made for them, and the optimized solutions are accumulated to be analyzed.

With this approach, we can take advantage of the resulting benefits in addition to the ones mentioned in Section I-C.

- Providing a direct solution: From the raw data of the devices and the market, the needed answers are provided.
- More accurate estimation of the failure risks: Failures are observed for a reasonable period of time, and they are not just *estimated* with different techniques. Thus a precise evaluation can be made, especially when a second-handed device is intended to be bought.
- Preparing a basis: This item is not directly related to the proposed approach; but when its prerequisite is fault isolation, a substructure is added to the work that can ameliorate other aspects of the optimization, such as prolonging preventive maintenance or decreasing the higher rate of failures at the initial times of installation due to improper installation or calibration (known as infant mortality in bathtub curve).

IV. Simulation and Results

To utilize an actual data set for the fault isolation part, the DAMADICS data set [21] is exploited. *General faults (F16-F18)* of this data set are determined precisely, but they are not adaptable directly for our purpose because of being imbalanced. Thus each class of faults and normal are separated and under-sampled to have comparable size; then they are split to train and test parts. The test part is utilized to test the classifier and to reproduce a suitable mock data set for 50 days, meeting our needs for a fictitious firm. The length of recorded data for each day is 10,000. And the distribution of faults for each device is brought in Table I.

TABLE I
Cost function parameters

	F16	**F17**	**F18**
Device 1	$\mu = 150, \sigma = 10$	$\mu = 80, \sigma = 5$	$\mu = 30, \sigma = 5$
Device 2	$\mu = 80, \sigma = 8$	$\mu = 40, \sigma = 5$	$\mu = 20, \sigma = 4$
Device 3	$\mu = 70, \sigma = 8$	$\mu = 32, \sigma = 5$	$\mu = 15, \sigma = 4$

Next, for designing a classifier, an ANN with the parameters mentioned in Table II is exploited that results in an acceptable classification as noted in Table III. Due to the relatively straightforward data, a simple network performs well. But it can be replaced with other networks or even other methods depending on the case.

TABLE II
ANN Parameters For Fault Isolation

Input Layer	**Hidden Layer**	**Output Layer**	**Loss**	**Optimizer**
Dim=7 (PCA Used)	Dim=50 Activ.: tanh	Dim=4 Activ.: sigmoid One-hot	mse	adamax $lr = 0.001$ $\beta_1 = 0.9$ $\beta_2 = 0.999$ $\epsilon = 1e-7$

TABLE III
Confusion Matrix of the Trained ANN Classifier

		Predict		
	N	**F16**	**F17**	**F18**
Actual	0.99603175	0.0026455	0	0.00132275
	0.14473684	0.81578947	0	0.03947368
	0	0	1	0
	0.16	0.14666667	0	0.69333333

For the market side, product prices produced by each of these 3 devices are different, and the price of buying the second-handed device is lower. The demand function of the product is exponentially decaying. For the cost of faults, a different price is assigned for each. Finally, the selling price of

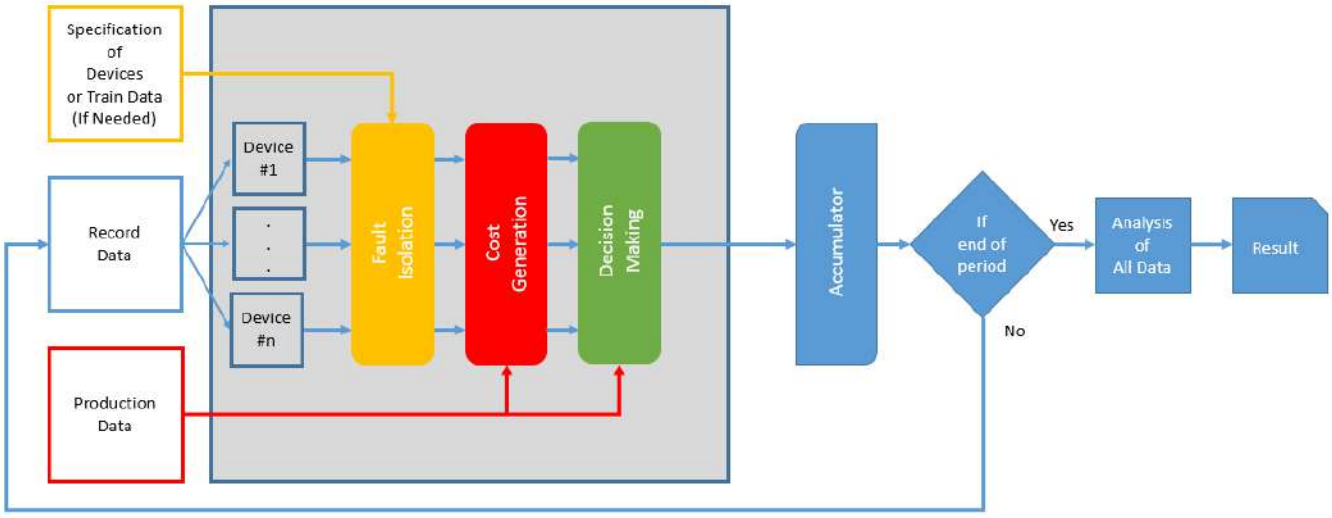


Fig. 1. The proposed procedure

the present device is fixed value as junk; but two other devices lose a percentage (10%) of their price at buying, then the price decays by passing the time. These are the main elements of the cost function that the stipulated parameters for this part are summarized in Table III.

TABLE IV
COST FUNCTION PARAMETERS

Days	Inflation(Daily)	Faults Quantity	F16 Cost	F17 Cost
50	0.9995	3	0.7	1.5
F18 Cost	Product Price by Device 1		Product Price by Device 2	
2.5	7.9		8.001	
Product Price by Device 3	Demand Function		Price of Device 2	
8.002	300exp(-0.003t)		40,000	
Price of Device 3	Sell Price of Device 1		Sell Price of Device 2	
70,000	7000		36000exp(-0.01t)	
Sell Price of Device 3				
63000exp(-0.017t)				

Finally, the simulation procedure is commenced by feeding data of each day to an ANN for fault isolation. Then the total cost of faults for each device is calculated, and with the market parameters, it is given to a GA (see Table V) for the minimizer of the cost. This is modeling the optimization as a zero-sum Stackelberg game.

TABLE V
GA MAIN PARAMETERS

Bounds	Strategy	Pop. Size	Mutation	CR
[(50,500), (1,3), (50,1000)]	best1bin	15	(0.5,1)	0.7

Returning to our main questions, mentioned in II:

- "When this device should be revised?"
- "Which device should it be replaced with?"
- "How long should the production line be kept after the replacement?"

first, the answer to this set of questions is provided for each day of the whole period that the results are illustrated in Table VI, Table VII, and Fig. 2.

TABLE VI
REMAINING DAYS UNTIL REPLACING THE PRESENT DEVICE

Up to Change	Range (days)		
	400-450	***450-500***	***Other***
Votes	3	44	3

TABLE VII
REMAINING DAYS UNTIL REPLACING THE PRESENT DEVICE

Next Device	Device Number	
	2	***3***
Votes	30	20

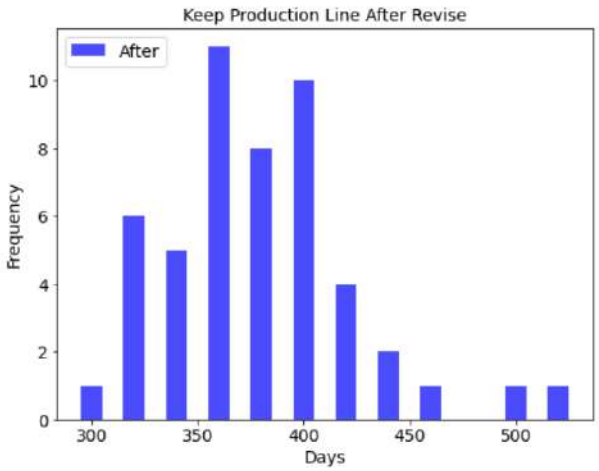


Fig. 2. Histogram of keeping the production line after replacing the device.

Next, these results will be our guide to determine the final answers.

- Revision Time:
  From Table VI, it can be concluded the best option for keeping the present degenerated device is in the range of 450 to 500 days (it must be noted that 500 is stipulated for the maximum viable option for the present device in the simulation). This means keeping the present device near its final days is the best option.
- The Next Device:
  Table VII indicates the next device to be bought is No.2, which is the second-handed one, but this decision is not as certain as the preceding one.
- Keeping Production Line Afterward:
  By inspecting Fig. 2, a range can be inferred the suitable option for keeping the production line after replacement lies between 350-450 days.

Finally, Fig. 3 provides additional information that reckons the profit of total production in today's value, that will be between 470K-500K dollars.

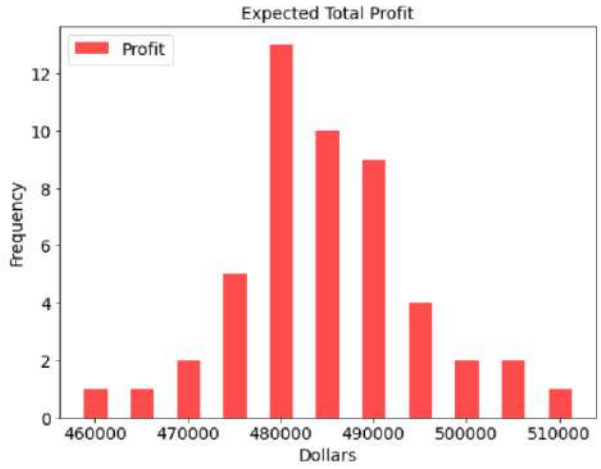


Fig. 3. Histogram of today's value of the expected profit after ending the production line.

It must be noted that except for this procedure, one could utilize the whole data of 50 days ensemble and minimize the cost function once. In this way, one particular answer will be found for each of the three questions, which is mathematically the optimal solution. But such an approach is not viable due to the underlying uncertainty of the modeling and the market. And providing a range is more preferred in real applications.

Finally, parameters selection for this procedure can be divided into three parts, concerning the main blocks in Fig. 1:

- Fault Isolation:
  Parameter selection for this part depends on the chosen method and the provided data set, and nothing general can be stated. Fortunately, the quality of this part can be measured and represented separately in a similar way as ours in Table III.
- Cost Generation:
  These parameters are directly related to the financial part, and they can be chosen arbitrarily. But for the example brought in this work, they are chosen meticulously (and unrealistic) to provide different possible scenarios and a proper illustration. And it must be noted that in such cases that the results are not unanimously agreeing on a selection, or even in some other cases seeming straightforward, the results are highly sensitive toward some market parameters. Thus the quality of the result depends on the appropriate estimation of the market.
- Decision Making:
  Similar to the Fault Isolation part, a wide variety of options are available for decision making. The method and its parameters must be selected with respect to the complexity of the search space, made by the cost function. Usually, PSO and GA can be handy options.

V. Conclusion and Future Work

In the present work, we proposed a solution to provide a set of decisions for device replacement. In order to meet a satisfactory result, we measured the deterioration level via a data-driven approach which is defining some faults. These faults can be isolated with the given sensory data automatically. We also indicated that modeling the aging effect can be performed by a broader sense of defining a fault. Finally, we made a fictitious process via a real fault isolation data set and simulated the proposed solution.

For future work, challenging each of explicit or implicit assumptions will lead the research to new levels that could increase the generality of the proposed solution in this paper, or introduce a better one. Additionally, readers are encouraged to explore other methods, such as fuzzy logic or ANFIS, and to consider possible amendments.

References

[1] S. Mathew and D. Kennedy, "A strategy for optimal equipment replacement", Production Planning and Control, vol. 14, no. 6, pp. 571-577, 2003.

[2] D. Weichert, P. Link, A. Stoll, S. Rüping, S. Ihlenfeldt and S. Wrobel, "A review of machine learning for the optimization of production processes", The International Journal of Advanced Manufacturing Technology, vol. 104, no. 5-8, pp. 1889-1902, 2019.

[3] G. Köksal, İ. Batmaz and M. Testik, "A review of data mining applications for quality improvement in manufacturing industry", Expert Systems with Applications, vol. 38, no. 10, pp. 13448-13467, 2011.

[4] H. Cheng and H. Chen, "Online parameter optimization in robotic force controlled assembly processes," 2014 IEEE International Conference on Robotics and Automation (ICRA), 2014, pp. 3465-3470.

[5] S. Genna, A. Simoncini, V. Tagliaferri and N. Uccia-rdello, "Optimization of the Sandblasting Process for a Better Electrodeposition of Copper Thin Films on Aluminum Substrate by Feedforward Neural Network", Procedia CIRP, vol. 62, pp. 435-439, 2017.

[6] Y. Rong, G. Zhang, Y. Chang and Y. Huang, "Integrated optimization model of laser brazing by extreme learning machine and genetic algorithm", The International Journal of Advanced Manufacturing Technology, vol. 87, no. 9-12, pp. 2943-2950, 2016.

[7] N. S. Shahraki and S. H. Zahiri, "Multi-objective optimization algorithms in analog active filter design," 2020 8th Iranian Joint Congress on Fuzzy and intelligent Systems (CFIS), 2020, pp. 105-109.
[8] H. Yektamoghadam and A. Nikoofard, "Fault Detection in Thermoelectric Energy Harvesting of Human Body", Journal of Electrical and Computer Engineering Innovations (JECEI), vol. 9, no. 1, pp. 57-66, 2021.
[9] M. Bazargan and J. Hartman, "Aircraft replacement strategy: Model and analysis", Journal of Air Transport Management, vol. 25, pp. 26-29, 2012.
[10] J. Hartman and C. Tan, "Equipment Replacement Analysis: A Literature Review and Directions for Future Research", The Engineering Economist, vol. 59, no. 2, pp. 136-153, 2014.
[11] Mahapatro, Arunanshu, and Pabitra Mohan Khilar. "Fault diagnosis in body sensor networks." International Journal of Computer Information Systems and Industrial Management Applications 5 (2012): 252-259.
[12] Sarcevic, Peter, Zoltan Kincses, and Szilveszter Pletl. "Online human movement classification using wrist-worn wireless sensors." Journal of Ambient Intelligence and Humanized Computing 10, no. 1 (2019): 89-106.
[13] Saneja, Bharti, and Rinkle Rani. "A scalable correlation-based approach for outlier detection in wireless body sensor networks." International Journal of Communication Systems 32, no. 7 (2019): e3918.
[14] Raheli, Bahare, Mohammad Taghi Aalami, Ahmed El-Shafie, Mohammad Ali Ghorbani, and Ravinesh C. Deo. "Uncertainty assessment of the multilayer perceptron (MLP) neural network model with implementation of the novel hybrid MLP-FFA method for prediction of biochemical oxygen demand and dissolved oxygen: a case study of Langat River." Environmental Earth Sciences 76, no. 14 (2017): 1-16.
[15] A. Shamshiri, M. Ghaznavi-Ghoushchi and A. Kariman, "ML-Based Aging Monitoring and Lifetime Prediction of IoT Devices with Cost-Effective Embedded Tags for Edge and Cloud Operability", IEEE Internet of Things Journal, pp. 1-1, 2021.
[16] Haghighi, Rouzbeh, Hossein Yektamoghadam, Majid Dehghani, and Amirhossein Nikoofard. "Generation Expansion Planning Using Game Theory Approach to Reduce Carbon Emission: A Case Study of Iran." Computers and Industrial Engineering (2021).
[17] Bresolin, Davide, and Marta Capiluppi. "A game-theoretic approach to fault diagnosis and identification of hybrid systems." Theoretical Computer Science 493 (2013): 15-29.
[18] Chen, X., Goh, M., Li, B., and Cheng, Y. (2021). Collection strategies and pricing decisions for dual channel EOL products. Computers and Industrial Engineering.
[19] S. Ding, Data-driven Design of Fault Diagnosis and Fault-tolerant Systems. London: Springer, 2014.
[20] S. Theodoridis, K. Koutroumbas and S. Theodoridis, Pattern recognition. Burlington, MA: Academic Press, 2009.
[21] M. Bartys, R. Patton, M. Syfert, S. de las Heras, and J. Quevedo, "Introduction to the damadics actuator fdi benchmark study", Control Engineering Practice, vol. 14, no. 6, pp. 577–596, June 2006.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

A Fractional Multi-commodity Network Flow Problem with Uncertain Multi-Choice Coefficients: Model Formulations and Solution Approach

1st Salim Bavandi, 2nd Seyed Hadi Nasseri and 3rd Chefi Triki

1st Department of Mathematics, *University of Mazandaran*, Babolsae, Iran, salimbavandi@yahoo.com
2nd Department of Mathematics, *University of Mazandaran*, Babolsar, Iran, nhadi57@gmail.com
3rd Kent Business School, *University of Kent*, Canterbury, United Kingdom, C.Triki@kent.ac.uk

***Abstract*— In this article, a Fractional Multi-Commodity Network Flow (FMCNF) problem under uncertainty conditions will be studied. due to increasing competition in the business world, which makes decision-makers deal with multiple options/ information for optimal decisions on a single task, we will look at multi-choice programming in a hybrid fuzzy random environment. The main objective of the present research is to provide the decision-maker with a model with the help of which he/she can manage unknown factors across a multi-commodity network. Given that this problem is herein investigated in a hybrid fuzzy-stochastic environment and includes multi-choice parameters, we used the probability–possibility approach, Lagrange interpolating polynomial, and the Charnes-Cooper variable transformation technique to convert the problem into a deterministic one. Finally, the efficiency of the proposed model will be evaluated by presenting a numerical instance.**

***Index Terms*— Multi-choice programming, fuzzy random variable, fractional programming, probability–possibility approach, multi-commodity network flow problem**

I. Introduction

In general, a network can transfer a variety of commodities, in which case it is called a Multi-Commodity Flow (MCNF) problem. In this network, the commodities are uniquely connected by using arcs. In this way, the arc capacity refers to the number of commodities that can pass through these arcs. Commodities may be distinguished by physical features or simply based on specific features. The MCNF problem has been widely used in the transportation industry, especially in the aviation industry, to synchronize crew and fleet assignments. It is worth noting that several objective functions are used in such cases, which include time, risk, cost, and environmental issues. Therefore, multi-objective flow models are more suitable for realworld models and their conditions compared to single-objective models. However, considering multiple objective functions in uncertainty conditions makes highly difficult the process of solving the MCNF problem with respect to traditional methods [1]. By considering only two objective functions, for example, minimizing costs and maximizing reliability or maximizing profits, it is possible to optimize the ratio of these objective functions. This type of problem, which is recognized as fractional programming, is a specific type of nonlinear programming problem with linear objective functions and constraints. Despite the presence of numerous deterministic models with thousands of variables and constraints in the industry, their results are oftentimes simply ignored by managers and decision-makers. This can be linked to the fact that such results are usually outcomes of deterministic modeling studies and become simply invalid under highly variable conditions in reality. Indeed, such changes in the condition may make previously optimal results either infeasible or suboptimal. On the other hand, many real-world situations come with multiple alternatives. For instance, a shipping carrier considers different shipping costs for transporting the same goods depending on the actual carrier (i.e., truck, ship, train, or aircraft). In the meantime, various companies have implemented different price discrimination strategies [17] based on price elasticity. For such companies, making optimal decisions when multiple choices are available is of paramount importance and imposes direct impacts on their profitability. During the last decades, the MCNF problem has become popular in the academic literature and an increasing number of researchers have become interested in it. It is a powerful operations research method for solving many complex problems, especially in the fields of transportation and communications [3]. The MCNF problem was introduced by Fords and Fulkerson [4] and Hu [5]. After that, many researchers tried to expand this model by adding different aspects to the problem. Even for a very simple MCNF problem with continuous flow and linear costs, it is very difficult

to find an optimal solution [6]. These difficulties are due to the complex nature of the constraint structure and the large number of variables that require extensive computational time [7]. Most research in this field has been done in fuzzy or stochastic conditions [8]–[10]. Among these, Ding [11] presented an MCNF problem with non-deterministic costs and capacities. He developed a decomposition-based algorithm to solve a (α, β)-minimum cost multi-commodity problem. A case of discrete cost multi-commodity flow with random demands has been investigated by Mejri et al. [12], where a penalty is considered for each un-routed demand. This problem requires finding a network topology that minimizes the total fixed start-up costs and expected penalties for not meeting the multi-commodity demand. In some cases, however, we deal with multiple alternative choices for each parameter of the problem, amongst which one should opt for proper choices to optimize the problem. The Multi-Choice Programming(MCP) was first introduced by Healy [13] in a study on a special case of mixed-integer programming. Since then, the wide application of the MCP in real-world decision-making processes motivated numerous researchers toward incorporating it into goal programming, transportation problems, the game theory, etc. [14]–[16]. As a common aspect among most of relevant research works, one may refer to deterministic nature of the alternative choices for the multi-choice parameters. Recently, however, Nasseri et al. [17] investigated a multi-choice linear programming problem in fuzzy-stochastic environment, wherein the alternative choices were represented by fuzzy-stochastic variables. On the other hand, when only two objective functions are considered, the ratio of these objective functions can be optimized. In fact, the problem of fractional programming can be solved by adopting some well-known approaches [18]. In addition, fractional programming occurs in many practical problems [19], [20] that can be followed in the other network analyses. So far, good research has been done on deficit planning under uncertainty. Recently, Nasseri et al. [21] have investigated the fuzzy random linear fractional programming problem. In their proposed model, the coefficients and scalars in the objective function are fuzzy numbers and technical coefficients and the right-hand values are the fuzzy random variables with a specified distribution. In another study, they offered a proposed approach to stochastic interval-valued linear fractional programming problems [22]. In this paper, we focus on a Multi-Commodity Network Flow (MCNF) problem with a fractional objective function in a hybrid fuzzy-stochastic environment, where the coefficients of the objective function in the nominator and denominator are multi-choice parameters represented by fuzzy-stochastic variables.

We generalize the deterministic fractional MCNF problem to its non-deterministic form. The rest of this paper is organized as follows. In Section 2, we briefly discuss some of the basic concepts and theorems related to the topic. In Section 3, we first present the deterministic model of the FMCNF and then generalize it to its non-deterministic form. The deterministic equivalent form of the FSFMCF problem is presented in Section 4. Finally, in Section 5, an example is provided to illustrate the efficiency of the model and algorithm.

II. Preliminaries

In this section, definitions and basic concepts related to probability theory and fuzzy set theory as well as fuzzy random variables will be presented as required.

Definition 1 (fuzzy variable): A fuzzy variable is a function of the possibility space to the set of real numbers.

Definition 2 (membership function): The membership function of a fuzzy variable $\tilde{A}$ is a function as $\mu_{\tilde{A}} : \mathcal{R} \to [0, 1]$ that assigns a degree of membership to each of the fuzzy values.

Definition 3: For a given α-cut, a α-cut set of the fuzzy set $\tilde{A}$ is an ordinary set of x elements such that the value of objective function is greater than α, i.e

$$A_\alpha = \{x | \mu_{\tilde{A}}(x) \geq \alpha\}$$

Definition 4: A fuzzy interval of type LR is displayed as $\tilde{A} : (\alpha, m_1, m_2, \beta)_{LR}$ where α and β are non-negative left and right spreads, respectively. Also, m_1 and m_2 are mean values of the fuzzy number $\tilde{A}$. The membership function of $\tilde{A}$ can be represented as following:

$$\mu_\xi(x) = \begin{cases} L\left(\frac{m_1 - x}{\alpha}\right), & \text{if } x \leq m_1 \\ 1, & \text{if } m_1 \leq x \leq m_2 \\ R, \left(\frac{x - m_2}{\beta}\right) & \text{if } x \geq m_2 \end{cases} \quad (1)$$

where, L and R are left and right continuous non-increasing functions from $[0, 1]$ to $[0, 1]$, respectively, such that $L(0) = R(0) = 1$ and $L(1) = R(1) = 0$. If $m_1 = m = m_2$, the fuzzy number LR is represented as following.

$$\mu_{\tilde{A}}(x) = \begin{cases} L\left(\frac{m - x}{\alpha}\right), & \text{if } x \leq m \\ R\left(\frac{x - m}{\beta}\right) & \text{if } x \geq m \end{cases} \quad (2)$$

which is called a triangular fuzzy number.

Definition 5: [23] The fuzzy random variable ξ is a function from the possibility space $(\Omega, \mathcal{A}, \Pr)$ to a set of fuzzy variables such that for every Borel set $\mathcal{B}$ of $\mathcal{R}$, $Pos\{\xi(\omega), \varepsilon \in \mathcal{B}\}$ is a measurable function of .

Theorem 1: [23]: Let ξ is a n-dimensional fuzzy random vector. If $f : \mathcal{R}^n \to \mathcal{R}$ is a measurable function, then $f(\xi)$ is a fuzzy random variable.

Definition 6: Let $\tilde{A}$ and $\tilde{B}$ are two independent fuzzy numbers with continuous membership functions. For the confidence level $\gamma \in [0, 1]$, we have:

$$Pos\left\{\tilde{A} \geq \tilde{B}\right\} \geq \gamma \Leftrightarrow \tilde{A}_\gamma^R \geq \tilde{B}_\gamma^L \quad (3)$$

Where $\tilde{A}_\gamma^R$ and $\tilde{B}_\gamma^L$ are the left and right extreme points of γ-cuts $\left[A_\gamma^L, A_\gamma^R\right]$ and $\left[B_\gamma^L, B_\gamma^R\right]$ of the fuzzy numbers $\tilde{A}$ and $\tilde{B}$.

III. Mathematical model

In this section, we first consider the deterministic form of the MCNF problem and then generalize it into a hybrid fuzzy stochastic environment. So, the directed network $G(N, A)$ is considered. In this network, the set of node is represented by $N = \{1, 2, ..., n\}$ and the set of arcs is indicated by $A = \{(i, j) | i, j \in N\}$. Suppose that K is a set of commodities $K = \{1, 2, ..., |K|\}$. c_{ijk} and d_{ijk} are the coefficients of cost arc (i, j). Other variables and parameters are as follows:
x_{ijk} is the amount of commodity of type k which crosses the

arc (i,j).
u_{ij} is an upper bound of the total flow for all commodities on the arc (i,j).
b_k is the amount of the commodity k which has to be sent from the origin s_k to the destination t_k.
Also assume that

$$b_i^k = \begin{cases} b_k & \text{if } i \text{ is the origin node of the commodity } k \\ -b_k & \text{if } i \text{ is the destination node of the commodity } k \\ 0 & \text{otherwise} \end{cases}$$

b_i^k is defined as the flow balance of commodity k on the node i. In fact, represent the supply or demand of commodity k on the node i. p and q are given constants.

In the crisp multi-commodity problem, the coefficient of the problem, are crisp non-negative values. In practice, however, these values are often not constant. In this situation, if enough data are available, the probability distribution of these data can be created using statistical analysis and inference based on the crisp data with appropriate measures. Unfortunately, in many cases, proper amounts of data are not available to determine the distribution. On the other hand, in some cases the collected data are inaccurate. So, the only hope in such circumstances is to use the inaccurate opinion of experts and professionals. In these cases, we face a fuzzy event among random variables. Under these conditions, let the decision-maker be dealt with multiple choices amongst which he/she must opt for only one. For this purpose, the decision-maker must follow an MCP approach.

In order to generalize the deterministic fractional multi-commodity flow problem (FMFP) to its uncertainty form, the parameters c_{ijk}, p, d_{ijk} and q are multi-choice parameters. Alternative values of c_{ijk} and p are represented by fuzzy-stochastic variables while those of d_{ijk} and q are fuzzy variables. Moreover, u_{ij} is herein a fuzzy-stochastic variable. As a result, an fractional multi-commodity network flow problem can be rewritten as a Fuzzy-Random Fractional Multi-Commodity Network Flow (FRFMCNF) problem through the following expression:

$$\begin{aligned} & Min\, f\left(x, \tilde{\tilde{c}}, \tilde{d}\right) = \frac{\sum_{k=1}^{K} \sum_{(i,j)\in A} \left\{\tilde{\tilde{c}}_{ijk}^{(1)}, \ldots, \tilde{\tilde{c}}_{ijk}^{(l_{ijk})}\right\} x_{ijk} + \left\{\tilde{\tilde{p}}^{(1)}, \ldots, \tilde{\tilde{p}}^{(k)}\right\}}{\sum_{k=1}^{K} \sum_{(i,j)\in A} \left\{\tilde{d}_{ijk}^{(1)}, \ldots, \tilde{d}_{ijk}^{(p_{ijk})}\right\} x_{ijk} + \left\{\tilde{q}^{(1)}, \ldots, \tilde{q}^{(l)}\right\}} \\ & s.t. \begin{cases} \sum\limits_{(i,j)\in A} x_{ijk} - \sum\limits_{(j,i)\in A} x_{jik} = b_i^k, \ \forall i \in N, \ \forall k \in K \\ \sum\limits_{k=1}^{K} x_{ijk} \le \tilde{\tilde{u}}_{ij}, \qquad \forall (i,j) \in A \\ x_{ijk} \ge 0, \qquad \forall (i,j) \in A, \forall k \in K \end{cases} \end{aligned} \tag{4}$$

where $\tilde{\tilde{c}}_{ijk}^{(r)} (r = 1, ..., k_{ij})$, $\tilde{\tilde{p}}^{(s)} (s = 1, ..., k)$, $\tilde{d}_{ijk}^{(t)} (t = 1, ..., l_{ij})$ and $\tilde{q}^{(u)} (u = 1, ..., l)$ are alternative choices c_{ijk}, p, d_{ijk} and q respectively. Given the existence of a fuzzy random variable in this objective function, we cannot calculate a specific value for every value of x. Therefore, its value cannot be determined by conventional methods. Also, the constraints in problem (4), do not specify a specific area due to the constraint of c_2'. In such cases, to deal with fuzzy random events, the probability-possibility approach with a predetermined confidence levels are utilized. Therefore, problem (4) can be rewritten as a Probability-Possibility Constrained Programming (PPCP) model as follows:

$$\begin{aligned} & Min\, f_0 \\ & s.t. \begin{cases} Pr\left[Pos\left[\frac{\sum_{k=1}^{K} \sum_{(i,j)\in A} \left\{\tilde{\tilde{c}}_{ijk}^{(1)}, \ldots, \tilde{\tilde{c}}_{ijk}^{(l_{ijk})}\right\} x_{ijk} + \left\{\tilde{\tilde{p}}^{(1)}, \ldots, \tilde{\tilde{p}}^{(k)}\right\}}{\sum_{k=1}^{K} \sum_{(i,j)\in A} \left\{\tilde{d}_{ijk}^{(1)}, \ldots, \tilde{d}_{ijk}^{(p_{ijk})}\right\} x_{ijk} + \left\{\tilde{q}^{(1)}, \ldots, \tilde{q}^{(l)}\right\}} \ge \delta\right] \ge \gamma \right] \\ \sum\limits_{(i,j)\in A} x_{ijk} - \sum\limits_{(j,i)\in A} x_{jik} = b_i^k, \quad \forall i \in N, \ \forall k \in K \\ Pr\left[Pos\left[\sum\limits_{k=1}^{K} x_{ijk} \le \tilde{\tilde{u}}_{ij}\right] \ge \theta\right] \ge \eta, \qquad \forall (i,j) \in A \\ x_{ijk} \ge 0, \qquad \forall (i,j) \in A, \ \forall k \in K \end{cases} \end{aligned} \tag{5}$$

In this model, the decision maker hopes to obtain the minimum value of f_0, so that $f\left(x, \tilde{\tilde{c}}, \tilde{d}\right)$ is equal to or less than f_0 with the confidence Levels of $\delta, \gamma \in [0,1]$.

The model presented in the previous section has a number of multi-choice parameters whose alternative values take the form of fuzzy-stochastic and fuzzy variables. This implies that the problem cannot be solved directly. In order to solve the problem, we begin by transforming the multi-choice parameters using a predefined interpolation polynomial. An interpolating polynomial introduces an integer variable for each multi-choice parameter. If the selected multi-choice parameter comes with k alternative values, the corresponding integer variable should include exactly k nodal points, with each node referring to exactly one functional value of the selected multi-choice parameter. Herein. The functional value of each node takes the form of a fuzzy stochastic and a fuzzy variable, and a multi-

choice parameter is substituted by an interpolating polynomial by using the Lagrange formula. Therefore, according to the Lagrange formula and interpolation polynomials for multi-choice parameters of the problem are obtained as follows:

$$
\begin{aligned}
&Min \;\; f_0 \\
&s.t. \quad \begin{cases}
Pr\left[Pos\left[\frac{\sum_{k=1}^{K}\sum_{(i,j)\in A} f_{\tilde{c}_{ijk}}(u_{ijk})x_{ijk}+f_{\tilde{p}}(w)}{\sum_{k=1}^{K}\sum_{(i,j)\in A} f_{\tilde{d}_{ijk}}(v_{ijk})x_{ijk}+f_{\tilde{q}}(z)} \le f_0\right] \ge \delta\right] \ge \gamma \\
\sum_{(i,j)\in A} x_{ijk} - \sum_{(j,i)\in A} x_{jik} = b_i^k, \quad \forall i \in N,\ \forall k \in K \\
Pr\left[Pos\left[\sum_{k=1}^{K} x_{ijk} \le \tilde{\tilde{u}}_{ij}\right] \ge \theta\right] \ge \eta, \quad \forall (i,j) \in A \\
x_{ijk} \ge 0, \qquad \forall (i,j) \in A,\ \forall k \in K \\
0 \le u_{ijk} \le k_{ij} - 1, \\
0 \le v_{ijk} \le l_{ij} - 1, \\
0 \le w \le k - 1, \\
0 \le z \le l - 1,
\end{cases}
\end{aligned} \tag{6}
$$

where $f_{\tilde{\tilde{c}}_{ijk}}(u_{ijk})$, $f_{\tilde{\tilde{p}}}(w)$, $f_{\tilde{d}_{ijk}}(v_{ijk})$ and $f_{\tilde{q}}(z)$ are interpolation polynomials for multi-choice parameters of the problem.

IV. Solution Approach

a solution method for problem (6) is to transform its constraints to their deterministic forms. Since the alternative choices $\tilde{\tilde{c}}_{ijk}^{r}(r = 1, ..., l_{ijk})$ and $\tilde{\tilde{p}}_s(s = 1, ..., k)$ as well as $\tilde{\tilde{u}}_{ij}$ are fuzzy stochastic parameters, one can write them in the following forms, respectively: $\tilde{\tilde{c}}_{ijk}^{r} = (c_{ijk}^{\alpha}, \bar{c}_{ijk}, c_{ijk}^{\beta})_{LR}$, $\tilde{\tilde{p}}_s = (p^{\alpha}, \bar{p}_s, p^{\beta})_{LR}$, and $\tilde{\tilde{u}}_{ij} = (u_{ij}^{\alpha}, \bar{u}_{ij}, u_{ij}^{\beta})_{LR}$, where the first and third elements represent the left and right tails, respectively, with the second elements being the center value. The second elements are normally distributed random variables with known mean values and variances. On the other hand, the variables $\tilde{d}_{ijk}^{t}(t = 1, ..., l_{ij})$ and $\tilde{q}_u(u = 1, ..., l)$ are fuzzy numbers. This implies that they can be expressed as $\tilde{d}_{ijk}^{t} = (d_{ijk}^{\alpha}, d_{ijk}, d_{ijk}^{\beta})_{LR}$ and $\tilde{q}_u = (q^{\alpha}, q, q^{\beta})$, respectively. Finally, according to definition 6 and chance-constrained approach and using the variable transformation technique introduced by Charness and Cooper [24] and Gupta and Jain [25], the crisp form of problem 4 will be as follows:

$$
\begin{aligned}
&Min \;\; \sum_{k=1}^{K}\sum_{(i,j)\in A}\left(f_{\tilde{\tilde{c}}_{ijk}}^{m}(u_{ijk}) - L^{-1}(\delta) f_{\tilde{\tilde{c}}_{ijk}}^{\alpha}(u_{ijk})\right) y_{ijk} + t f_{\tilde{\tilde{p}}}^{m}(w) \\
&\quad + tL^{-1}(\delta) f_{\tilde{\tilde{p}}}^{\alpha}(w) + \Phi^{-1}(\gamma) \sqrt{\sum_{k=1}^{K}\sum_{(i,j)\in A} \sigma^2_{f^{m}_{\tilde{\tilde{c}}_{ijk}}(u_{ijk})} y_{ijk}^2 + t^2 \sigma^2_{f^{m}_{\tilde{\tilde{p}}}(w)}} \\
&s.t. \quad \begin{cases}
\sum_{k=1}^{K}\sum_{(i,j)\in A}\left(f_{\tilde{d}_{ijk}}^{m}(v_{ijk}) + R^{-1}(\delta) f_{\tilde{d}_{ijk}}^{\beta}(v_{ijk})\right) y_{ijk} + t f_{\tilde{q}}^{m}(z) + tR^{-1}(\delta) f_{\tilde{q}}^{\beta}(z) = 1 \\
\sum_{(i,j)\in A} y_{ijk} - \sum_{(j,i)\in A} y_{jik} = b_i^k t, \qquad \forall i \in N,\ \forall k \in K \\
\sum_{k=1}^{K} y_{ijk} \le \left(u_{ij}^{m} + R^{-1}(\theta) u_{ij}^{\beta} + \Phi^{-1}(1-\eta)\sigma_{u_{ij}^{m}}\right) t,\ \forall (i,j) \in A \\
y_{ijk}, t \ge 0, \\
0 \le u_{ijk} \le k_{ij} - 1, \\
0 \le v_{ijk} \le l_{ij} - 1, \\
0 \le w \le k - 1, \\
0 \le z \le l - 1,
\end{cases}
\end{aligned} \tag{7}
$$

The objective function of this problem is convex according to $\Phi^{-1}(\gamma) \ge 0$ for every $\gamma \in [0.5, 1)$ and its constraints are linear. Therefore, problem 7 is a convex problem which by solving it, we will be able to provide a solution to problem 4.

V. An illustrative example

[26] Consider a network consisting of 6 nodes and 10 edges that is shown by a directed network in Figure1. Data related to the problem are presented in Tables I and II where $p = 0$ and $\tilde{q} = \left\{\tilde{q}^{(1)} = (0.5, 0.3), \tilde{q}^{(2)} = (1, 0.4), \tilde{q}^{(3)} = (1.5, 0.5)\right\}$. In this example, the coefficients of the nominator of the fraction and the edge capacity are fuzzy random coefficients and the coefficients of the denominator of the fraction are assumed to be fuzzy numbers. Each of the fuzzy and fuzzy random coefficients has a symmetrical triangular membership function. Hence, we show them as pairs of $(a(\omega), \alpha)$ and (m, α).

Using the data of Tables I and II, the deterministic model of the problem is obtained in the form of problem (7). This model is coded in Mathematica 10 software for different γ and η

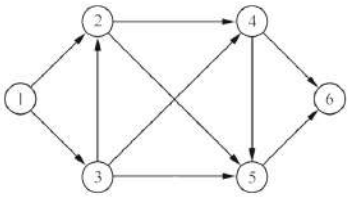


Fig. 1. Schematic representation of the network of the Example1

TABLE I
THE NUMBER OF DEMANDS BETWEEN ORIGIN AND DESTINATIONS

Commodity k	$O(k)$	$D(k)$	Demand b^k
1	1	4	1
2	1	5	3
3	1	6	2

TABLE II
THE RELATED DATA OF THE EXAMPLE1

$Edge$	k	$\tilde{\tilde{c}}_{ijk}^{(1)}$	$\tilde{\tilde{c}}_{ijk}^{(2)}$	$\tilde{\tilde{c}}_{ijk}^{(3)}$	$\tilde{d}_{ijk}^{(1)}$	$\tilde{d}_{ijk}^{(2)}$	$\tilde{d}_{ijk}^{(3)}$	$\tilde{\tilde{u}}_{ij}$
$(1,2)$	1	$(N(1.5,0.25),1)$	$(N(2,0.5),1.5)$	$(N(2.5,0.75),2)$	$(1,2)$	$(1.5,1)$	$(2,2)$	$(N(9,0.25),4)$
	2	$(N(1.5,0.5),2)$	$(N(2,0.25),2.5)$	$(N(2.5,1),3)$	$(1,1)$	$(1.5,1)$	$(2,1)$	
	3	$(N(1,0.25),1)$	$(N(1.5,0.5),1.5)$	$(N(1,0.75),2)$	$(1.5,2)$	$(2,3)$	$(3,3)$	
$(1,3)$	1	$(N(11,0.5),3)$	$(N(11.5,0.75),3.5)$	$(N(12,0.25),4)$	$(3,2)$	$(4,1)$	$(5,3)$	$(N(3,0.5),2)$
	2	$(N(12,1),5)$	$(N(12.5,0.75),5.5)$	$(N(13,0.5),6)$	$(3,1)$	$(4,2)$	$(5,2)$	
	3	$(N(13,1),7)$	$(N(13.5,0.75),7.5)$	$(N(14,1),8)$	$(3,2)$	$(4,1)$	$(5,2)$	
$(3,2)$	1	$(N(2,0.5),3)$	$(N(2.5,0.25),3.5)$	$(N(3,0.75),4)$	$(1,5)$	$(2,6)$	$(3,7)$	$(N(1.5,0.25),3)$
	2	$(N(1.5,1),2)$	$(N(2,0.5),2.5)$	$(N(2.5,0.25),3)$	$(1,6)$	$(2,5)$	$(2,6)$	
	3	$(N(1.5,0.25),3)$	$(N(2,0.5),3.5)$	$(N(2.5,0.75),4)$	$(1,2)$	$(2,3)$	$(3,4)$	
$(2,4)$	1	$(N(2,1),3)$	$(N(2.5,0.5),3.5)$	$(N(3,1),4)$	$(1,5)$	$(2,4)$	$(3,5)$	$(N(1.5,0.5),1)$
	2	$(N(1,0.25),1)$	$(N(1.5,0.25),1.5)$	$(N(2,0.5),2)$	$(2,1)$	$(3,2)$	$(4,3)$	
	3	$(N(3.5,0.25),6)$	$(N(4,0.5),6.5)$	$(N(4.5,0.25),7)$	$(1,1)$	$(2,3)$	$(3,4)$	
$(2,5)$	1	$(N(5,1),5)$	$(N(5.5,0.5),5.5)$	$(N(6,1),6)$	$(4,3)$	$(5,2)$	$(6,3)$	$(N(1.5,0.75),5)$
	2	$(N(2,0.25),2)$	$(N(2.5,0.5),2.5)$	$(N(3,0.5),3)$	$(5,2)$	$(6,1)$	$(7,2)$	
	3	$(N(3.5,0.5),4)$	$(N(4,1),4.5)$	$(N(4.5,0.5),5)$	$(4,3)$	$(5,1)$	$(6,5)$	
$(3,4)$	1	$(N(8,1),7)$	$(N(8.5,0.5),7.5)$	$(N(9,0.25),8)$	$(6,1)$	$(7,2)$	$(8,3)$	$(N(5,0.5),3)$
	2	$(N(7,1),5)$	$(N(7.5,0.5),5.5)$	$(N(8,0.5),6)$	$(6,2)$	$(7,1)$	$(8,2)$	
	3	$(N(5.5,0.5),2)$	$(N(6,0.5),2.5)$	$(N(6.5,1),3)$	$(5,1)$	$(6,2)$	$(7,3)$	
$(3,5)$	1	$(N(15,1),6)$	$(N(15.5,0.5),6.5)$	$(N(16,0.5),7)$	$(4,2)$	$(5,1)$	$(6,1)$	$(N(2,1),4)$
	2	$(N(16,1),8)$	$(N(16.5,0.75),8.5)$	$(N(17,0.5),9)$	$(4,1)$	$(5,2)$	$(6,3)$	
	3	$(N(13,0.5),3)$	$(N(13.5,0.5),3.5)$	$(N(14,1),4)$	$(1,5)$	$(2,6)$	$(3,7)$	
$(4,5)$	1	$(N(12,0.25),5)$	$(N(12.5,1),5.5)$	$(N(13,0.25),6)$	$(1,1)$	$(2,1)$	$(3,2)$	$(N(1.5,0.5),1)$
	2	$(N(11,1),3)$	$(N(11.5,0.5),3.5)$	$(N(12,1),4)$	$(1,2)$	$(2,3)$	$(3,1)$	
	3	$(N(11,0.5),4)$	$(N(11.5,0.75),4.5)$	$(N(12,0.5),5)$	$(1,1)$	$(2,3)$	$(3,2)$	
$(4,6)$	1	$(N(1,0.5),1)$	$(N(1.5,0.25),1.5)$	$(N(2,0.5),2)$	$(1,2)$	$(2,2)$	$(3,2)$	$(N(4,5,0.5),7)$
	2	$(N(1.5,0.5),2)$	$(N(2,0.5),2.5)$	$(N(2.5,0.25),3)$	$(1,4)$	$(2,5)$	$(3,5)$	
	3	$(N(3,0.5),4)$	$(N(4,0.5),4.5)$	$(N(4.5,1),5)$	$(1,2)$	$(2,3)$	$(3,3)$	
$(5,6)$	1	$(N(3,0.5),4)$	$(N(3.5,1),4.5)$	$(N(4,0.75),5)$	$(4,3)$	$(5,4)$	$(6,4)$	$(N(2.5,0.5),1)$
	2	$(N(3.5,1),3)$	$(N(4,0.5),3.5)$	$(N(4.5,0.75),4)$	$(5,1)$	$(6,3)$	$(7,2)$	
	3	$(N(4.5,1),6)$	$(N(5,0.5),6.5)$	$(N(6,1),7)$	$(4,4)$	$(5,3)$	$(6,5)$	

values and the obtained solution is presented in both numerical and schematic forms in Table III and Figure 2 respectively.

The results obtained from this example show that the proposed model can be very efficient if it has suitable inputs. As you can see, for different probability levels, different but close solutions are obtained. On the other hand, in this model we make decisions that ensure its feasibility as much as possible. In fact, constraints are not violated except in emergencies and unpredictability. It is worth noting that here only an idea was presented for the selection of and but in practice, this value is determined by the decision maker in a way that meets the safety and security needs.

REFERENCES

[1] M. Ghatee, S.M. Hashemi, "Some concepts of the fuzzy multi-commodity flow problem and their application in fuzzy network design," Math. Comput. Model. vol. 49, pp. 1030–1043, 2009.

[2] H. R. Varian, "Price discrimination," Handb Ind Organ. vol. 1, pp. 597–654, 1989.

[3] S. Bavandi, S. H. Nasseri, C. Triki, Optimal Decision Making in Fuzzy Stochastic Hybrid Uncertainty Environments and Their Application in Transportation Problems, In: Cao B. (eds) Fuzzy Information and Engineering-2019. Advances in Intelligent Systems and Computing,1094. Springer, Singapore, 2020.

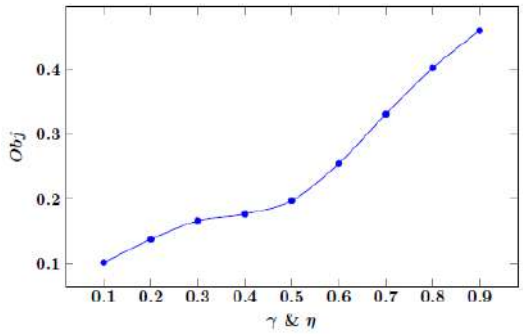


Fig. 2. The optimal value of the objective function for different values γ and η.

TABLE III
OPTIMAL SOLUTION FOR THE EXAMPLE.

Edge	(Commodity 1)	Commodity 2	Commodity 3
$(1,2)$	1.00	3.00	2.00
$(1,3)$	0.00	0.00	0.00
$(3,2)$	0.00	0.00	0.00
$(2,4)$	1.00	0.00	1.64
$(2,5)$	0.00	3.00	0.36
$(3,4)$	0.00	0.00	0.00
$(3,5)$	0.00	0.00	0.00
$(4,5)$	0.00	0.00	0.00
$(4,6)$	0.00	0.00	1.64
$(5,6)$	0.00	0.00	0.36

[4] L. Jr. Ford, D. Fulkerson, "A suggested computation for maximal multi-commodity network flows," Manag Sci. vol. 5, pp. 97–101, 1958.

[5] T. C. Hu, "Multi-commodity network flows," Oper Res. vol. 11, pp. 344–360, 1963.

[6] P. Mahey, M. C. de Souza, "Local optimality conditions for multi-commodity flow problems with separable piecewise convex costs," Oper Res Lett. vol. 35, pp. 221–226, 2007.

[7] H. Nagamochi, T. Ibaraki, "Multi-commodity flows in certain planar directed networks," Discrete Appl Math. vol. 27, pp. 125–145, 1990.

[8] P.M. Bialon, "A randomized rounding approach to a k-splittable multi-commodity flow problem with lower path flow bounds affording solution quality guarantees," Telecommun Syst. vol. 64, pp. 525–542, 2017.

[9] Q. Liu, Q. Zhao, and W. Zang, "Study on multi-objective optimization of flow allocation in a multi-commodity stochastic-flow network with unreliable nodes," J. Appl. Math. Comput. vol. 28, pp. 185–198, 2008.

[10] S. Sekar, L. Zheng, L. J. Ratliff, and B. Zhang "Uncertainty in Multi-commodity Routing Networks: When Does It Help?," in IEEE Transactions on Automatic Control. vol. 65, pp. 4600–4615, 2020.

[11] S. Ding, "Uncertain minimum cost multi-commodity flow problem," Soft Comput. vol. 21, pp. 223–231, 2017.

[12] I. Mejri, S. B. Layeb, M. Haouari, and F. Z. Mansour, "A simulation-optimization approach for the stochastic discrete cost multi-commodity flow problem," Engineering Optimization. vol. 52, pp. 507-826 2019.

[13] Jr. WC. Healy, "Multiple choice programming (a procedure for linear programming with zero-one variables)," Oper Res. vol. 12, pp. 122–138, 1964.

[14] A. Charnes, W. W.Cooper, "Management models and industrial applications of linear programming," Manage Sci. vol. 4, pp. 38–91, 1957.

[15] D. Dutta, S. Murthy, "Multi-choice goal programming approach for a fuzzy transportation problem," Int J Res Rev Appl Sci. vol. 2, pp. 132–139, 2010.

[16] A. Panda, C. B. Das, "Multi-choice linear programming for matrix game," Appl Math Comput. vol. 237, pp. 411–418, 2014.

[17] S. H. Nasseri, S. Bavandi, "Multi-choice Linear Programming in Fuzzy Random Hybrid Uncertainty Environment and Their Application in Multi-commodity Transportation Problem," Fuzzy Information and Engineering. vol. 12, pp. 109–122, 2021.

[18] T. C. Hu, "Multi-commodity network flows," Oper Res. vol. 11, pp. 344–360, 1963.

[19] Q. Liu, Q. Zhao, W. Zang. "Study on multi-objective optimization of flow allocation in a multi-commodity stochastic-flow network with unreliable nodes," J. Appl. Math. Comput. vol. 28, pp. 185–198, 2008.

[20] Y. K. Liu, B. Liu, "Fuzzy random variables: a scalar expected value operator," Fuzzy Optimization and Decision Making. vol. 2, pp. 143–160, 2003.

[21] P. Mahey, M. C. de Souza, "Local optimality conditions for multi-commodity flow problems with separable piecewise convex costs," Oper Res Lett. vol. 35, pp. 221–226, 2007.

[22] R. D. McBride, "advances in solving the multi-commodity-flow problem," Interfaces. vol. 28, pp. 32–41, 1998.

[23] Y. K. Liu, B. Liu, "Fuzzy random variables: a scalar expected value operator," Fuzzy Optimization and Decision Making. vol. 2, pp. 143–160, 2003.

[24] A. Charnes, W. W. Cooper, "Programming with linear functional," Naval Research Logistics Quarterly, vol. 9, pp. 181–186, 1962.

[25] S.N. Gupta, R. K. Jain, "Stochastic fractional programming under chance constraints with random technology matrix," Acta Ciencia Indica. vol. 3, pp. 191–198, 1986.

[26] R.K. Ahuja, Magnanti T.L. and Orlin J.B. Network Flows: Theory, Algorithms, and Applications, Prentice Hall, Upper Saddle River, New Jersey, 1993.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

A New Approach For Solving The Interval Type-2 Fuzzy Transportation Problem Based On interval Linear Programming Problems

1st Moslem Javanmard, 2nd Hassan Mishmast Nehi

1st Math teacher with a PhD, *Education Department*, Zabol, Iran, javanmard.moslem@gmail.com
2nd Mathematics Department, *University of Sistan and Baluchestan*, Zahedan, Iran, hmnehi@hamoon.ac.ir

***Abstract*— In real world problems, the parameters of the solid transportation problems (supply, demand and capacity of vehicles) are not always accurate. Therefore, in dealing with these problems due to the existence of uncertainty and ambiguity in the data should use fuzzy optimization. Hence, in the main part of fuzzy optimization, we need to model and solve fuzzy programming problems that its parameters are interval type-2 fuzzy numbers. While, the studies on model of interval type-2 fuzzy linear programming problem have been very limited. So, in this paper, we consider the solid transportation problem that its parameters are interval type-2 triangular fuzzy numbers. Then, we proposed a new approach to solve it by extension the concept of the nearest interval approximation. Finally, in special case, we explain and analyze the numerical example of interval type-2 fuzzy solid transportation problem.**

***Index Terms*— best-worst cases; interval linear programming; interval type-2 fuzzy linear programming; solid transportation problem; nearest interval approximation**

I. Introduction

So far, many methods such as, ranking or ordering methods and approximation techniques [3, 4, 7, 8] to solve the fuzzy linear programming problem are proposed, that each methods has its own advantages and disadvantages. One of the applications of the every existing defuzzification methods is to solve single or multi objective type-1 and interval type-2 fuzzy linear programming problem. In the general interval type-2 fuzzy linear programming problem, all coefficients and numbers used at it, are interval type-2 fuzzy numbers. The solid transportation problems are considered a special category of interval linear programming problems. In the fuzzy solid transportation problem a product from source m to destination n will be sent with rth-vehicle. The solid transportation problem with type-1 fuzzy parameters has been discussed by many of researchers [9, 11, 13]. However, until now, the limited researches about transportation problem with type-2 fuzzy parameters have been done. Kundu et. al, [10, 12], the solid transportation problem modeled and solved with general type-2 fuzzy parameters. Also, Figuera and Hernandez [2], for first time, presented the solid transportation problem with interval type-2 fuzzy parameters. In this paper, we consider the solid transportation problem that its parameters (supply, demand and capacity of vehicles) are interval type-2 triangular fuzzy numbers. Then, we suggested a new approach to solve it, by extension the nearest interval approximation and using the best-worst cases (BWC) method for all created interval linear programming problems. Eventually, in simple case, we solve and explain the numerical example of interval type-2 fuzzy solid transportation problem.

II. Overview of Methodology

Consider the general fuzzy solid transportation problem that the parameters of supply, demand and capacity of vehicle are interval type-2 triangular fuzzy numbers, as

$$\begin{aligned} \text{Min} \quad & Z = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{r} c_{ijk}x_{ijk} \\ \text{s.t.} \quad & \\ & \sum_{j=1}^{n}\sum_{k=1}^{r} x_{ijk} \le \tilde{a}_i, \quad i = 1,2,...,m, \\ & \sum_{i=1}^{m}\sum_{k=1}^{r} x_{ijk} \ge \tilde{b}_j, \quad j = 1,2,...,n, \\ & \sum_{i=1}^{m}\sum_{j=1}^{n} x_{ijk} \le \tilde{e}_k, \quad k = 1,2,...,r, \\ & x_{ijk} \ge 0, \ \forall i,j,k, \end{aligned} \tag{1}$$

where, the coefficient c_{ijk} (the definite value), is the penalty for carrying an unit of product from source i to destination j with the k-th vehicle. This penalty may be include the cost of delivery time, delivery safety or warranty. Also, respectively, the supply, demand and vehicle capacity values as $\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_m$, $\tilde{b}_1, \tilde{b}_2, ..., \tilde{b}n$ and $\tilde{e}_1, \tilde{e}_2, ..., \tilde{e}_r$ are that for each i, j and

k represented as the following interval type-2 triangular fuzzy numbers:

$$\tilde{a}_i = (a_i^l, a_i^u) = \Big((a_{i1}^l, a_{i2}, a_{i3}^l), (a_{i1}^u, a_{i2}, a_{i3}^u)\Big)$$
$$\tilde{b}_j = (b_j^l, b_j^u) = \Big((b_{j1}^l, b_{j2}, b_{j3}^l), (b_{j1}^u, b_{j2}, b_{j3}^u)\Big)$$
$$\tilde{e}_k = (e_k^l, e_k^u) = \Big((e_{k1}^l, e_{k2}, e_{k3}^l), (e_{k1}^u, e_{k2}, e_{k3}^u)\Big).$$

Since, an interval type-2 triangular fuzzy number, is composed the infinite union of the characteristic type-1 triangular fuzzy numbers, we use the traditional embedded type-1 triangular fuzzy numbers for all the interval type-2 triangular fuzzy numbers of problem (1). Therefore, with the help of the nearest interval approximation of them, we convert the interval type-2 triangular fuzzy linear programming problem (1) into the following interval linear programming problem that the beginning and end of each of which its intervals are intervals (namely: IILP problem (2)), as:

$$\begin{aligned}
&\text{Min} \quad z^e = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{r} c_{ijk}x_{ijk} \\
&\text{s.t.} \\
&\quad \sum_{j=1}^{n}\sum_{k=1}^{r} x_{ijk} \le [\frac{\check{a}_i + a_{i2}}{2}, \frac{\hat{a}_i + a_{i2}}{2}], \quad i = 1, 2, ..., m, \\
&\quad \sum_{i=1}^{m}\sum_{k=1}^{r} x_{ijk} \ge [\frac{\check{b}_j + b_{j2}}{2}, \frac{\hat{b}_j + b_{j2}}{2}], \quad j = 1, 2, ..., n, \\
&\quad \sum_{i=1}^{m}\sum_{j=1}^{n} x_{ijk} \le [\frac{\check{e}_k + e_{k2}}{2}, \frac{\hat{e}_k + e_{k2}}{2}], \quad k = 1, 2, ..., r, \\
&\quad x_{ijk} \ge 0, \ \forall i, j, k,
\end{aligned} \tag{2}$$

where, $\hat{a}_i \in [a_{i3}^l, a_{i3}^u], \check{a}_i \in [a_{i1}^u, a_{i1}^l]$ $\hat{b}_j \in [b_{j3}^l, b_{j3}^u], \check{b}_j \in [b_{j1}^u, b_{j1}^l]$ $\hat{e}_k \in [e_{k3}^l, e_{k3}^u], \check{e}_k \in [e_{k1}^u, e_{k1}^l]$.

By the BWC method [1] for IILP problem (2), the best and the worst modeles are obtained as follows

Best model:

$$\begin{aligned}
&\text{Min} \quad z^e = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{r} c_{ijk}x_{ijk} \\
&\text{s.t.} \\
&\quad \sum_{j=1}^{n}\sum_{k=1}^{r} x_{ijk} \le \frac{\hat{a}_i + a_{i2}}{2}, \quad i = 1, 2, ..., m, \\
&\quad \sum_{i=1}^{m}\sum_{k=1}^{r} x_{ijk} \ge \frac{\check{b}_j + b_{j2}}{2}, \quad j = 1, 2, ..., n, \\
&\quad \sum_{i=1}^{m}\sum_{j=1}^{n} x_{ijk} \le \frac{\hat{e}_k + e_{k2}}{2}, \quad k = 1, 2, ..., r, \\
&\quad x_{ijk} \ge 0, \ \forall i, j, k,
\end{aligned} \tag{3}$$

where, $\hat{a}_i \in [a_{i3}^l, a_{i3}^u]$ $\check{b}_j \in [b_{j1}^u, b_{j1}^l]$ and $\hat{e}_k \in [e_{i3}^l, e_{i3}^u]$.

Worst model:

$$\begin{aligned}
&\text{Min} \quad z^e = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{r} c_{ijk}x_{ijk} \\
&\text{s.t.} \\
&\quad \sum_{j=1}^{n}\sum_{k=1}^{r} x_{ijk} \le \frac{\check{a}_i + a_{i2}}{2}, \quad i = 1, 2, ..., m, \\
&\quad \sum_{i=1}^{m}\sum_{k=1}^{r} x_{ijk} \ge \frac{\hat{b}_j + b_{j2}}{2}, \quad j = 1, 2, ..., n, \\
&\quad \sum_{i=1}^{m}\sum_{j=1}^{n} x_{ijk} \le \frac{\check{e}_k + e_{k2}}{2}, \quad k = 1, 2, ..., r, \\
&\quad x_{ijk} \ge 0, \ \forall i, j, k,
\end{aligned} \tag{4}$$

where, $\check{a}_i \in [a_{i1}^u, a_{i1}^l]$ $\hat{b}_j \in [b_{j3}^l, b_{j3}^u]$ and $\check{e}_k \in [e_{i1}^u, e_{i1}^l]$.

It is clear, the problems (3) and (4) are interval linear programming models (namely: ILP problem (3) and (4)). Hence, we obtain theirs sub-models by reusing the BWC method [1] for them. The best-best and the best-worst models obtained from (3), which are listed below

1) The best-best model

$$\begin{aligned}
&\text{Min} \quad z^e = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{r} c_{ijk}x_{ijk} \\
&\text{s.t.} \\
&\quad \sum_{j=1}^{n}\sum_{k=1}^{r} x_{ijk} \le \frac{a_{i3}^u + a_{i2}}{2}, \quad i = 1, 2, ..., m, \\
&\quad \sum_{i=1}^{m}\sum_{k=1}^{r} x_{ijk} \ge \frac{b_{j1}^u + b_{j2}}{2}, \quad j = 1, 2, ..., n, \\
&\quad \sum_{i=1}^{m}\sum_{j=1}^{n} x_{ijk} \le \frac{e_{k3}^u + e_{k2}}{2}, \quad k = 1, 2, ..., r, \\
&\quad x_{ijk} \ge 0, \ \forall i, j, k.
\end{aligned} \tag{5}$$

2) The best-worst model

$$\begin{aligned}
&\text{Min} \quad z^e = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{r} c_{ijk}x_{ijk} \\
&\text{s.t.} \\
&\quad \sum_{j=1}^{n}\sum_{k=1}^{r} x_{ijk} \le \frac{a_{i3}^l + a_{i2}}{2}, \quad i = 1, 2, ..., m, \\
&\quad \sum_{i=1}^{m}\sum_{k=1}^{r} x_{ijk} \ge \frac{b_{j1}^l + b_{j2}}{2}, \quad j = 1, 2, ..., n, \\
&\quad \sum_{i=1}^{m}\sum_{j=1}^{n} x_{ijk} \le \frac{e_{k3}^l + e_{k2}}{2}, \quad k = 1, 2, ..., r, \\
&\quad x_{ijk} \ge 0, \ \forall i, j, k.
\end{aligned} \tag{6}$$

Also, the worst-best and the worst-worst models obtained from (4), which are listed below

3) The worst-best model

$$\begin{aligned}
&\text{Min} \quad z^e = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{r} c_{ijk}x_{ijk} \\
&\text{s.t.} \\
&\qquad \sum_{j=1}^{n}\sum_{k=1}^{r} x_{ijk} \le \frac{a^l_{i1}+a_{i2}}{2}, \quad i=1,2,...,m, \\
&\qquad \sum_{i=1}^{m}\sum_{k=1}^{r} x_{ijk} \ge \frac{b^l_{j3}+b_{j2}}{2}, \quad j=1,2,...,n, \\
&\qquad \sum_{i=1}^{m}\sum_{j=1}^{n} x_{ijk} \le \frac{e^l_{k1}+e_{k2}}{2}, \quad k=1,2,...,r, \\
&\qquad x_{ijk} \ge 0, \ \forall i,j,k.
\end{aligned} \tag{7}$$

4) The worst-worst model

$$\begin{aligned}
&\text{Min} \quad z^e = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{r} c_{ijk}x_{ijk} \\
&\text{s.t.} \\
&\qquad \sum_{j=1}^{n}\sum_{k=1}^{r} x_{ijk} \le \frac{a^u_{i1}+a_{i2}}{2}, \quad i=1,2,...,m, \\
&\qquad \sum_{i=1}^{m}\sum_{k=1}^{r} x_{ijk} \ge \frac{b^u_{j3}+b_{j2}}{2}, \quad j=1,2,...,n, \\
&\qquad \sum_{i=1}^{m}\sum_{j=1}^{n} x_{ijk} \le \frac{e^u_{k1}+e_{k2}}{2}, \quad k=1,2,...,r, \\
&\qquad x_{ijk} \ge 0, \ \forall i,j,k.
\end{aligned} \tag{8}$$

As well as, by placing the core of the interval type-2 triangular fuzzy numbers [5, 6] with themselves in the problem (1), we find the fifth linear programming model namely the middle model (9), as follows

5- The middel model

$$\begin{aligned}
&\text{Min} \quad Z = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{r} c_{ijk}x_{ijk} \\
&\text{s.t.} \\
&\qquad \sum_{j=1}^{n}\sum_{k=1}^{r} x_{ijk} \le a_{i2}, \quad i=1,2,...,m, \\
&\qquad \sum_{i=1}^{m}\sum_{k=1}^{r} x_{ijk} \ge b_{j2}, \quad j=1,2,...,n, \\
&\qquad \sum_{i=1}^{m}\sum_{j=1}^{n} x_{ijk} \le e_{k2}, \quad k=1,2,...,r, \\
&\qquad x_{ijk} \ge 0, \ \forall i,j,k
\end{aligned} \tag{9}$$

Eventually, by solving these five (the best-best (5), best-worst (6), worst-best (7), worst-worst (8) and middel (9)) models, we get the answer to the interval type-2 triangular fuzzy linear programming problem (1), as:

$$\tilde{z}^{opt} = \Big((\underline{\hat{z}}^{opt}, \bar{z}^{opt}, \underline{\tilde{z}}^{opt}), (\overline{\hat{z}}^{opt}, \bar{z}^{opt}, \overline{\tilde{z}}^{opt}) \Big).$$

In next, we will solve the introduced transportation problem in a simple case with two sources (supply or inventory), two destinations (demand) and two vehicles (capacity level).

III. Practical Example: fuzzy solid transportation Problem in special case

Consider the solid transportation problem whose parameters are supply, demand and the capacity of the vehicles are interval type-2 triangular fuzzy numbers, as follows:

$$\begin{aligned}
&\text{Min} \quad Z = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} c_{ijk}x_{ijk} \\
&\text{s.t.} \\
&\qquad \sum_{j=1}^{2}\sum_{k=1}^{2} x_{ijk} \le \tilde{a}_i, \quad i=1,2, \\
&\qquad \sum_{i=1}^{2}\sum_{k=1}^{2} x_{ijk} \ge \tilde{b}_j, \quad j=1,2, \\
&\qquad \sum_{i=1}^{2}\sum_{j=1}^{2} x_{ijk} \le \tilde{e}_k, \quad k=1,2, \\
&\qquad x_{ijk} \ge 0, \ \forall i,j,k \in \{1,2\},
\end{aligned} \tag{10}$$

where, the coefficient c_{ijk} (the definite value), is the penalty for carrying an unit of product from source i to destination j with the k-th vehicle. Also, two sources (supply or inventory), two destinations (demand) and two vehicles (capacity level) are the parameters of the problem, which in the form of linguistic variables according to the table (I), by the decision maker in the table (II) introduced.

TABLE I
LINGUISTIC VARIABLES REPRESENT THE PARAMETERS OF INTERVAL TYPE-2 FUZZY TRANSPORTATION PROBLEM

Input parameters	Linguistic conditions
Inventory: **Low** $\le \tilde{3}2,\ \tilde{3}2 <$ **Appropriate** $\le \tilde{4}2,\ \tilde{4}2 <$ **Abundant**	
Demand: **A little** $\le \tilde{3}0,\ \tilde{3}0 <$ **medium** $\le \tilde{3}3,\ \tilde{3}3 <$ **much**	
Vehicle capacity: **Insignificant** $\le \tilde{3}1,\ \tilde{3}1 <$ **Limited** $\le \tilde{4}3,\ \tilde{4}3 <$ **Top**	

According to the decision maker, the parameters of the solid transportation problem are introduced as interval type-2

triangular fuzzy numbers, which are given below:

$$\tilde{a}_1 = \Big((a^l_{11}, a_{12}, a^l_{13}), (a^u_{11}, a_{12}, a^u_{13})\Big) = \Big((39, 42, 44), (38, 42, 45)\Big),$$

$$\tilde{a}_2 = \Big((a^l_{21}, a_{22}, a^l_{23}), (a^u_{21}, a_{22}, a^u_{23})\Big) = \Big((30, 33, 35), (29, 33, 36)\Big),$$

$$\tilde{b}_1 = \Big((b^l_{11}, b_{12}, b^l_{13}), (b^u_{11}, b_{12}, b^u_{13})\Big) = \Big((29, 29.5, 31), (28, 29.5, 32)\Big),$$

$$\tilde{b}_2 = \Big((b^l_{21}, b_{22}, b^l_{23}), (b^u_{21}, b_{22}, b^u_{23})\Big) = \Big((32, 32.5, 35.75), (31, 32.5, 36.25)\Big),$$

$$\tilde{e}_1 = \Big((e^l_{11}, e_{12}, e^l_{13}), (e^u_{11}, e_{12}, e^u_{13})\Big) = \Big((30, 32, 34.5), (29.5, 32, 37)\Big),$$

$$\tilde{e}_2 = \Big((e^l_{21}, e_{22}, e^l_{23}), (e^u_{21}, e_{22}, e^u_{23})\Big) = \Big((41, 43, 46), (39, 43, 47)\Big).$$

Also, the selected coefficients c_{ij1} and c_{ij2} for each $i = 1, 2$ and $j = 1, 2$, by the decision maker are available, as the following:

$$c_{111} = 10,\ c_{121} = 8,\ c_{211} = 13,\ c_{221} = 9,$$

$$c_{112} = 11,\ c_{122} = 14,\ c_{212} = 15,\ c_{222} = 12$$

TABLE II
TRIANGULAR-LINGUISTIC OF PARAMETERS SELECTED BY THE DECISION MAKER AS (LOWER MEMBERSHIP FUNCTION, UPPER MEMBERSHIP FUNCTION)

problem parameters	triangular-linguistic parameters
$\tilde{a}_1$	(Appropriate, Abundant)
$\tilde{a}_2$	(Low, Appropriate)
$\tilde{b}_1$	(A little , A little)
$\tilde{b}_2$	(medium, Much)
$\tilde{e}_1$	(Insignificant, Limited)
$\tilde{e}_2$	(Top , Top)

So, by inserting the existing assumptions, the problem (10) becomes as follows:

$$\begin{aligned}
&\text{Min } Z = 10x_{111} + 8x_{121} + 13x_{211} + 9x_{221} + \\
&\qquad 11x_{112} + 14x_{122} + 15x_{212} + 12x_{222} \\
&\text{s.t.} \\
&\sum_{j=1}^{2}\sum_{k=1}^{2} x_{1jk} \le \Big((39, 42, 44), (38, 42, 45)\Big), \\
&\sum_{j=1}^{2}\sum_{k=1}^{2} x_{2jk} \le \Big((30, 33, 35), (29, 33, 36)\Big), \\
&\sum_{i=1}^{2}\sum_{k=1}^{2} x_{i1k} \ge \Big((29, 29.5, 31), (28, 29.5, 32)\Big), \\
&\sum_{i=1}^{2}\sum_{k=1}^{2} x_{i2k} \ge \Big((32, 32.5, 35.75), (31, 32.5, 36.25)\Big), \\
&\sum_{i=1}^{2}\sum_{j=1}^{2} x_{ij1} \le \Big((30, 32, 34.5), (29.5, 32, 37)\Big), \\
&\sum_{i=1}^{2}\sum_{j=1}^{2} x_{ij2} \le \Big((41, 43, 46), (39, 43, 47)\Big), \\
&\qquad x_{ijk} \ge 0,\ \forall i, j, k \in \{1, 2\}
\end{aligned} \tag{11}$$

Also, to obtain the optimal value of the problem (11), by applying the approach based on the nearest interval approximation [5, 6], we solve the best-best (5), best-worst (6) , worst-best (7), worst-worst (8) and the middle (9) models, respectively, as follow:

1- The best-best model

$$\begin{aligned}
&\text{Min } \ Z = 10x_{111} + 8x_{121} + 13x_{211} + 9x_{221} + \\
&\qquad 11x_{112} + 14x_{122} + 15x_{212} + 12x_{222} \\
&\text{s.t.} \\
&x_{111} + x_{121} + x_{112} + x_{122} \le 43.5, \\
&x_{211} + x_{221} + x_{212} + x_{222} \le 34.5, \\
&x_{111} + x_{112} + x_{211} + x_{212} \ge 28.75, \\
&x_{121} + x_{122} + x_{221} + x_{222} \ge 31.75, \\
&x_{111} + x_{211} + x_{121} + x_{221} \le 34.5, \\
&x_{112} + x_{212} + x_{122} + x_{222} \le 45, \\
&\qquad x_{ijk} \ge 0,\ \forall i, j, k \in \{1, 2\}.
\end{aligned} \tag{12}$$

2- The best-worst model

$$\begin{aligned}\text{Min}\quad & Z = 10x_{111} + 8x_{121} + 13x_{211} + 9x_{221} + \\ & 11x_{112} + 14x_{122} + 15x_{212} + 12x_{222}\end{aligned}$$

s.t.

$$\begin{aligned}
&x_{111} + x_{121} + x_{112} + x_{122} \leq 43,\\
&x_{211} + x_{221} + x_{212} + x_{222} \leq 34,\\
&x_{111} + x_{112} + x_{211} + x_{212} \geq 29.25,\\
&x_{121} + x_{122} + x_{221} + x_{222} \geq 32.25,\\
&x_{111} + x_{211} + x_{121} + x_{221} \leq 33.25,\\
&x_{112} + x_{212} + x_{122} + x_{222} \leq 44.5,\\
&x_{ijk} \geq 0,\ \forall i,j,k \in \{1,2\}.
\end{aligned} \tag{13}$$

3- The worst-best model

$$\begin{aligned}\text{Min}\quad & Z = 10x_{111} + 8x_{121} + 13x_{211} + 9x_{221} \\ & + 11x_{112} + 14x_{122} + 15x_{212} + 12x_{222}\end{aligned}$$

s.t.

$$\begin{aligned}
&x_{111} + x_{121} + x_{112} + x_{122} \leq 43.5,\\
&x_{211} + x_{221} + x_{212} + x_{222} \leq 34.5,\\
&x_{111} + x_{112} + x_{211} + x_{212} \geq 28.75,\\
&x_{121} + x_{122} + x_{221} + x_{222} \geq 31.75,\\
&x_{111} + x_{211} + x_{121} + x_{221} \leq 34.5,\\
&x_{112} + x_{212} + x_{122} + x_{222} \leq 45,\\
&x_{ijk} \geq 0,\ \forall i,j,k \in \{1,2\}.
\end{aligned} \tag{14}$$

4- The worst-worst model

$$\begin{aligned}\text{Min}\quad & Z = 10x_{111} + 8x_{121} + 13x_{211} + 9x_{221} \\ & + 11x_{112} + 14x_{122} + 15x_{212} + 12x_{222}\end{aligned}$$

s.t.

$$\begin{aligned}
&x_{111} + x_{121} + x_{112} + x_{122} \leq 40,\\
&x_{211} + x_{221} + x_{212} + x_{222} \leq 31,\\
&x_{111} + x_{112} + x_{211} + x_{212} \geq 30.75,\\
&x_{121} + x_{122} + x_{221} + x_{222} \geq 34.375,\\
&x_{111} + x_{211} + x_{121} + x_{221} \leq 30.75,\\
&x_{112} + x_{212} + x_{122} + x_{222} \leq 41,\\
&x_{ijk} \geq 0,\ \forall i,j,k \in \{1,2\}.
\end{aligned} \tag{15}$$

5- The middel model

$$\begin{aligned}\text{Min}\quad & Z = 10x_{111} + 8x_{121} + 13x_{211} + 9x_{221} + \\ & 11x_{112} + 14x_{122} + 15x_{212} + 12x_{222}\end{aligned}$$

s.t.

$$\begin{aligned}
&x_{111} + x_{121} + x_{112} + x_{122} \leq 42,\\
&x_{211} + x_{221} + x_{212} + x_{222} \leq 33,\\
&x_{111} + x_{112} + x_{211} + x_{212} \geq 29.5,\\
&x_{121} + x_{122} + x_{221} + x_{222} \geq 32.5,\\
&x_{111} + x_{211} + x_{121} + x_{221} \leq 32,\\
&x_{112} + x_{212} + x_{122} + x_{222} \leq 43,\\
&x_{ijk} \geq 0,\ \forall i,j,k \in \{1,2\}.
\end{aligned} \tag{16}$$

The result of solving the best-best (5), best-worst (6) , worst-best (7), worst-worst (8) and the middle (9) models are given in table (III),

By collecting the obtained results from the table (III), the optimal value of the objective function of the interval type-2 fuzzy solid transportation problem (11) as an interval type-2 triangular fuzzy number in the table (IV) provided.

TABLE III
RESULTS OF THE SUB-MODELS REQUIRED TO SOLVE THE INTERVAL TYPE-2 TRIANGULAR FUZZY SOLID TRANSPORTATION PROBLEM

The name of the model	The optimal value
The best-best model	$\overline{\hat{z}}^{opt} = 584.5$
The best-worst model	$\underline{\hat{z}}^{opt} = 597.25$
The worst-best model	$\underline{\bar{z}}^{opt} = 639$
The worst-worst model	$\overline{\bar{z}}^{opt} = 649.25$
The middel model	$\bar{z} = 606$

TABLE IV
THE OPTIMAL VALUE AS AN INTERVAL TYPE-2 TRIANGULAR FUZZY NUMBER

interval type-2 triangular fuzzy number	The optimal value
$\Big((597.25, 606, 639), (584.5, 606, 649.25)\Big)$	$\tilde{z}^{opt}$

IV. CONCLUSION

Until now, many methods such as, ranking or ordering methods and approximation techniques to solve the fuzzy linear programming problem are proposed, that each methods has its own advantages and disadvantages and it would be a hard task to decide which one is the best choice. Thus each time the defuzzification method should be chosen to the particular problem, that is the responsibility of the decision maker. In this paper, the main problem as the solid transportation problem considered that its parameters (supply, demand and capacity of vehicles) are interval type-2 triangular fuzzy numbers. Since, each interval type-2 triangular fuzzy numbers, the main problem is composed the infinite union of the characteristic type-1 triangular fuzzy numbers. Therefore, by placing approximate closed intervals with each interval type-2 triangular fuzzy numbers in the main problem, it converts to an interval-interval linear programming (IILP) problem. Also, twice using BWC method on IILP problem and placing core of interval type-2 triangular fuzzy numbers, the main problem becomes five alternative linear programming models. It is clear that the new employed approach is better than the ranking methods, because the use of closed intervals is a generalization of the placement of crisp numbers in the main problem. Infact, here, less information are lost than other methods. So, in here, an easy and useful way to solve the fuzzy problem with interval type-2 triangular fuzzy numbers is provided.

REFERENCES

[1] M. Allahdadi, H. Mishmast Nehi, H. A. Ashayerinasab, M. Javanmard, (2016), Improving the modified interval linear programming method by new techniques, Information Sciences, 339, pp. 224–236.

[2] J. C. Figueroa, G. Hernandez, A transportation model with interval type-2 fuzzy demands and supplies, Lecture Notes in Computer Science, 7389 (2012) 610–617.

[3] M. Javanmard, H. Mishmast Nehi, Rankings and operations for interval type-2 fuzzy numbers: a review and some new methods, Journal of Applied Mathematics and Computing, DOI: https://doi.org/10.1007/s12190-018-1193-9, First Online: 31 May 2018.

[4] M. Javanmard, H. Mishmast Nehi, Arithmetic operations and ranking methods for interval type-2 fuzzy numbers; A review and some new methods, 9th International Iranian Operation Research conference, IORC, Shiraz, Iran (2016), 928-933.

[5] M. Javanmard, H. Mishmast Nehi, Solving the interval type-2 fuzzy linear programming problem by the nearest interval approximation, 6th Iranian Joint Congress on Fuzzy and Intelligent Systems, CFIS, Kerman, Iran, DOI: https://doi.org/10.1109/CFIS.2018.8336609, Publisher IEEE, Feb 2018.

[6] M. Javanmard, H. Mishmast Nehi, A Solving Method for Fuzzy Linear Programming Problem with Interval Type-2 Fuzzy Numbers, Int. J. Fuzzy Syst., DOI: https://doi.org/10.1007/s40815-018-0591-3, Published Online: 17 January 2019.

[7] M. Javanmard, H. Mishmast Nehi, Interval type-2 triangular fuzzy numbers; New ranking method and evaluation of some reasonable properties on it, 5th Iranian Joint Congress on Fuzzy and Intelligent Systems, CFIS, 8003587 (2017), 4–6.

[8] M. Javanmard, H. Mishmast Nehi, Solving interval type-2 fuzzy linear programming problem with a new ranking function method, 5th Iranian Joint Congress on Fuzzy and Intelligent Systems, CFIS, 8003586 (2017), 1–3.

[9] F. Jimnez, J.L. Verdegay, Solving fuzzy solid transportation problems by an evolutionary algorithm based parametric approach, European Journal of Operational Research, 117 (1999) 485–510.

[10] P. Kundu, S. Kar, M. Maiti, Fixed charge transportation problem with type-2 fuzzy variables, Information Sciences, 255 (2014) 170–186.

[11] P. Kundu, S. Kar, M. Maiti, Multi-objective multi-item solid transportation problem in fuzzy environment, Applied Mathematical Modeling, 37 (2012) 2028–2038.

[12] P. Kundu, S. Kar, M. Maiti, S. Majumder, A method to solve linear programming problem with interval type-2 fuzzy parameters, Fuzzy Optim Decis Making, DOI: https://doi.org/10.1007/s10700-018-9287-2, Published Online: 6 Feb 2018.

[13] L. Yang, L. Liu, Fuzzy fixed charge solid transportation problem and algorithm, Applied Soft Computing, 7 (2007) 879–889.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

A new contrast enhancement method for Color dark and low-light images

Meysam Alavi[1], Mehrdad Kargari[2]

[1] *Information Technology dept., Tarbiat Modares, Tehran, Iran*, meysamalavi@modares.ac.ir
[2] *Information Technology dept, Tarbiat Modares, Tehran, Iran*, M_kargari@modares.ac.ir

***Abstract*— Image contrast enhancement is a preprocessing phase that improves the performance of image processing applications such as pattern recognition and computer vision. Many images have poor quality due to low- luminance and low-contrast, which must be changed before further processing. It is significant for medical imaging because the low-light intensity makes it challenging to diagnose and analyze specific diseases accurately. In addition, when the depth information of a low-light image is unknown, the drawback of illumination enhancement becomes very challenging. Since Contrast enhancement of low-light images with non-uniform illumination is complex, it may lead to inefficient contrast enhancement methods. In other words, in such cases, if contrast enhancement methods are used to increase the contrast of dark areas, the bright regions become over-enhanced, which may lead to the disappearance of the details of these areas. Given this problem, in this paper, a new method is proposed to increase the image contrast that can improve the contrast in both areas of the image without adversely affecting the details of the image. The result represented the efficiency of proposed methods compared to other image enhancement methods based on full-reference and no-reference metrics such as PSNR, MSE, NIQE, and BRISQUE.**

***Index Terms* — Image processing, image enhancement, contrast enhancement, low-light images.**

I. INTRODUCTION

Image enhancement plays a substantial role in image processing and analysis. Improving the image includes increasing contrast, increasing the edge, reducing noise, which enhances the perception of the human supervisor of the information in the image[1]. Contrast enhancement is considered a commonly used image enhancement method. Contrast is defined as segregation in intensity level between an image's bright and dark regions[2]. Problems like device malfunction, the inexperience of taking photographs, and imaging with low or high illuminance intensity, may lead to reduced contrast. As a result, degraded images may disturb the visual aesthetics[3].

Insufficient contrast in images reduces the ability of the observer to understand and analyze the image[4]. It is significant for medical imaging because the low-light intensity makes it challenging to diagnose and analyze specific diseases accurately[1]. Furthermore, in dark images with non-uniform illumination, low contrast leads to the disappearance of the details in both the dark region and the bright region. In such cases, contrast enhancement methods may lead to over-enhanced in the bright area. Hence some details of the bright region are deleted[5]. Dark or weak illumination image enhancement is one of the challenging research areas which have applications in night vision, forensic applications, and underwater enhancement[6]. Owing to the significance of this issue, different methods have been proposed by researchers. Although these methods have tried to improve the contrast, some cases have problems such as loss of image detail, lack of preservation of structural features of the image, and increased noise. Given the importance of the issue, this paper a method for improving contrast based on the separation of red, green, and blue channels of an image and improving each channel, will present that while improving image contrast and noise reduction, the structural features of the image are preserved to an acceptable level. The results demonstrate the efficiency of the proposed method compared to other methods.

The organization of this paper is as follows: In Section 2, an overview of some techniques of image contrast enhancement is addressed. The proposed method describes in Section 3. Section

4 presents experimental results. Finally, in Section 5, the conclusion and future work are drawn.

II. LITERATURE REVIEW

In recent years, researchers have proposed different methods of contrast enhancement, which provide better images for various applications of image processing. Generally, these methods are categorized into global, local, and hybrid contrast enhancement[7]. On the other hand, these methods are categorized into direct and indirect methods [8]. Most of the methods used in the previous research belong to the second group. One of the most favorite techniques is histogram equalization(HE), which can be global or local. The aim of HE is the intensity values distribution in an image such that the lower contrast regions can achieve higher contrast [9]. Paleness in some areas, loss of some details, and increased noise in some image areas are the disadvantages of this method [10]. Contrast Limited Adaptive Histogram Equalization (CLAHE) to overcome the problems of the HE is presented, which is a local HE method[11]. CLAHE provides maximum entropy while flattening the histogram; therefore, it is more suitable for medical images, especially for X-Ray images, where the low contrast is a challenge. [9]. In the following years, based on the Histogram equalization, various methods were proposed, which can be referred to as BBHE[12], DSIHE[13], RMSHE[14], RSIHE[15], MMBEBHE[16], BPDHE[17]. In[18], Chiu et al. proposed an enhancement method based on the adaptive gamma correction method for low-light images, which used weight distribution to correct the pixel intensity of the input image. This method faced over-enhancement in bright regions of the image. Also, in [19], a technique based on gamma correction to increase the brightness level of a low-contrast image and, Consequently, correct the probability distribution of bright pixels is presented. In [20] a robust Retinex model is presented that adds noise terms to handle low-light image enhancement in the case of drastic noise. Results of the research show this method has good affective in low-light image enhancement. Therefore, it can be generalized to enhance images in underwater, remote sensing, and hazy or dusty conditions. In [21], a fusion-based method to improve low-light images is proposed that can handle various low-light parameters in the same, which images can be enhanced in different weak illumination conditions such as non-uniform illumination, backlighting, and nighttime. Shi et al. proposed an image enhancement method that used a dual-channel method to enhance nighttime low illumination images. The results demonstrate substantial validity of the approach both visually and quantitatively compared to other available methods[22]. Regarding the literature review, this paper presents a new method to enhance the image, which we are able to enhance the contrast in both regions of the image without changing the main points. It also preserves brightness and structural image features while noise is reduced.

III. THE PROPOSED METHOD

In this section, the proposed method is presented. In this method, First, the original image in the RGB-color space is divided into three red, green, and blue channels. Then the algorithm is applied to each RGB channel separately. For each RGB channel of the input image, fuzzification is done using the formula:

$$\mu(g_{ij}) = \frac{g - g_{\min}}{g_{\max} - g_{\min}} \tag{1}$$

Where g is the intensity of the image ranges from 0 to L-1. Minimum and maximum values of the intensity for each RGB channel are illustrated as $g_{\min}$ and $g_{\max}$. The new membership function has been demonstrated with μ_{new}, which is composed of two parts: natural fuzzy complement and Sugeno. These a mentioned in Equations (2) and (3), respectively[23].

$$\bar{\mu} = 1 - \mu \tag{2}$$

$$C_{\lambda}(\mu) = \frac{1-\mu}{1+\lambda\mu} \tag{3}$$

μ_{new}, Based on Equations (2), (3), and (4), Eq. (5) can be concluded.

$$\mu_{new} = 1 - C_{\lambda}(\mu) \tag{4}$$

$$\mu_{new} = \frac{\mu(1+\lambda_{Opt})}{1+\lambda_{Opt}\mu} \tag{5}$$

In the proposed method, an index composed of three metrics, Feature similarity Index (FSIM), structural similarity index (SSIM), and Absolute mean brightness error (AMBE), is defined to improve the contrast of an image while preserving brightness and structural image features. This index is shown in Eq. (6).

$$\text{Evaluation Index} = \frac{FSIM * SSIM}{AMBE} \tag{6}$$

Using SSIM, the structure of the input image with the enhanced image is compared[24], calculated by Eq. (7).

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{7}$$

Where μ_x, μ_y are the mean intensity of the original and enhanced image, respectively, and σ_x^2, σ_y^2 and σ_{xy} are the variances and the covariance, respectively. A higher value of SSIM shows better performance[25]. The Maximum SSIM value is 1. The FSIM is used as a qualitative metric for the enhanced images. It is calculated as follows:

$$FSIM = \frac{\sum S_L * PC_m}{\sum PC_m} \tag{8}$$

PC And *S* are the maximum local adaptation and the similarity matrix, respectively[26]. The AMBE is calculated as below:

$$AMBE = |M(X) - M(Y)| \quad (9)$$

The low value of AMBE illustrates that the average brightness of the enhanced image is close to or equal to the average brightness of the input image[27]. Changing λ leads to creating images with different brightness levels. In other words, changing λ from -1 to infinity creates completely dark to completely white images, and consequently, one of the values in the range of -1 to infinity leads to the optimal state of image contrast. This value of λ is indicated optimal λ and is defined as λ_{opt} . After increasing the λ value, the image is created based on the μ_{new} . Then the evaluation index is calculated and compared with the threshold. If the indicator is higher than the defined threshold, the λ value increases again. This process continues, and when it stops, the last λ value is selected as the optimal λ. Finally, the output image is created based on the optimal λ_{opt} value using Eq. (5). Based on the experimental criterion, the threshold value is considered 0.05, which has the best result. The mentioned process is performed separately for each red, green, and blue channel, and finally, the outputs are merged, and image contrast is enhanced. The pseudo-code of the proposed method is shown in Algorithm 1.

Algorithm 1

Input: Color Image
Output : Enhanced Image
For R,G,B channel of an image
Start
Step1 : the input image is fuzzified using Eq.(1)
Start loop
Step 2 : Increase λ and Calculate μ_{new} using Eq.(5) .
Step 4: Calculate $\frac{FSIM * SSIM}{AMBE}$
if $\frac{FSIM * SSIM}{AMBE} <$ threshold
Go to Step 2
End loop
Step 4 : Determine optimal λ_{opt}
Step 5 : Calculate μ_{new} based on λ_{opt}
Step 6 : Image Defuzzification
End
Step 7: Merge of outputs

A. *Evaluation metrics*

Given the close relationship between image evaluation and human visual perception, finding a universal standard for quantifying the quality of an improved image is difficult[21]. In general, image quality assessment (IQA) is categorized into no reference and full reference. To evaluate the quantitative analysis of the proposed method with other methods are considered, different metrics like Entropy, PSNR, MSE, NIQE[28], and BRISQUE[29].

Entropy is defined as in Eq. (10).

$$Entropy = -\sum_{i=0}^{i-1} p(i)\log_2 p(i) \quad (10)$$

According to entropy theory, the more extensive and detailed the information in an image, the greater the entropy (IE) [30]. The PSNR of an image is used for measuring the image denoising effect[31], as shown in Eq. (11). In which L is 256 for 8-bit gray-scale images and MSE is the mean squares error, as shown in Eq. (12). According to Eq. (11), the higher the PSNR value, the lower the mean square error, so the better the image quality[25].

$$PSNR = 10 * \log_{10}\left(\frac{(L-1)^2}{MSE}\right) \quad (11)$$

$$MSE(X,Y) = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}|X(i,j) - Y(i,j)|^2}{M \times N} \quad (12)$$

BRISQUE is considered as the no-reference image quality category metric, which compares input images to a default model computed from images of natural scenes with similar distortions[32]. The smaller score of NIQE and BRISQUE illustrates better perceptual quality [33].

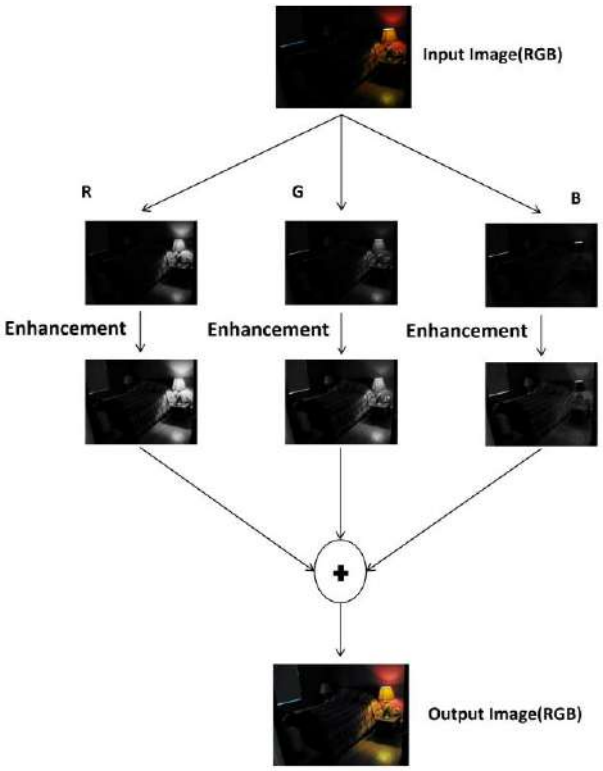


Figure 1. Outline of the proposed method

IV. EXPERIMENTAL RESULTS

In this research, the Exdark dataset is used for the experimental analysis. The dataset is available on GitHub [34]. To the assessment of the proposed method, we used 45 images, which included dark and low-light images. Also, the proposed method is compared with HE, method of [21], method of[20], AGC[3]. The comparison of average results for images based on evaluation metrics (no reference and full reference) is shown in Figures 2 to 7.

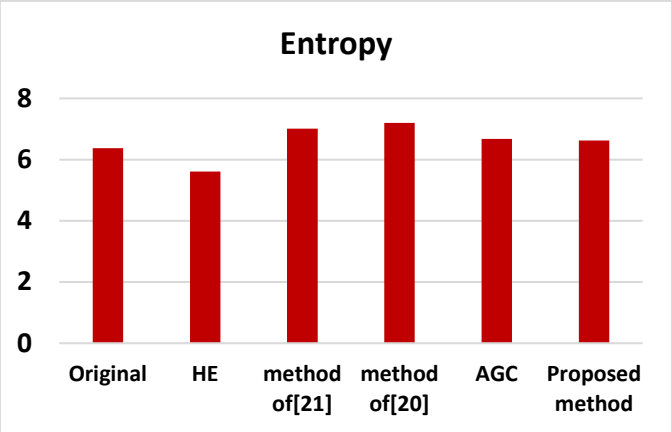


Figure 2. Entropy average for total images in a different method

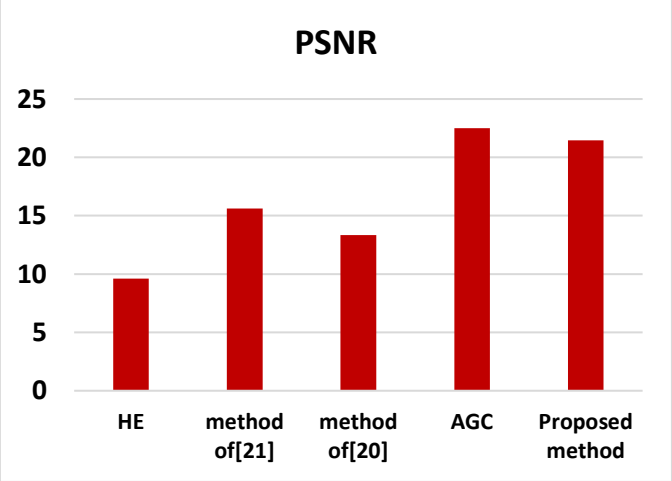


Figure 3. PSNR average for total images in a different method

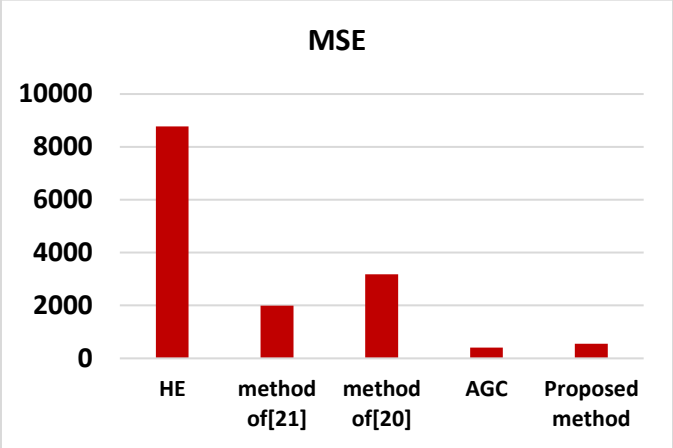


Figure 4. MSE average for total images in a different method

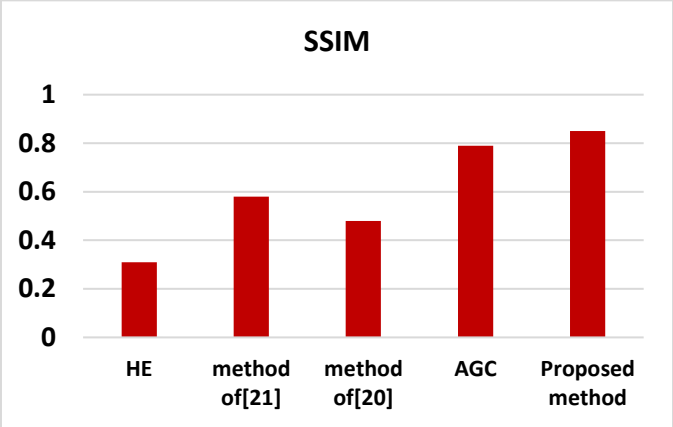


Figure 5. SSIM average for total images in a different method

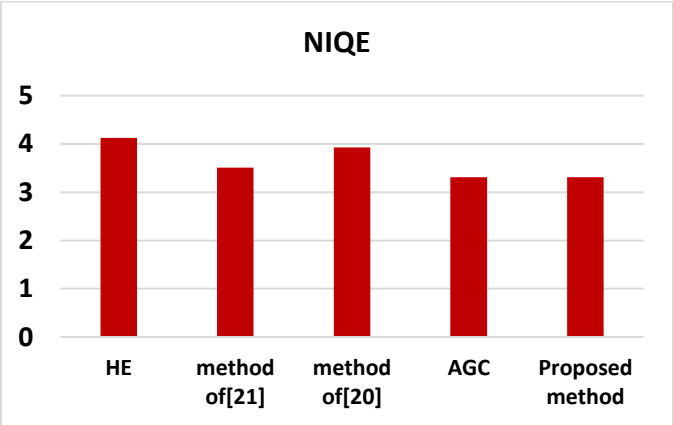


Figure 6. NIQE average for total images in a different method

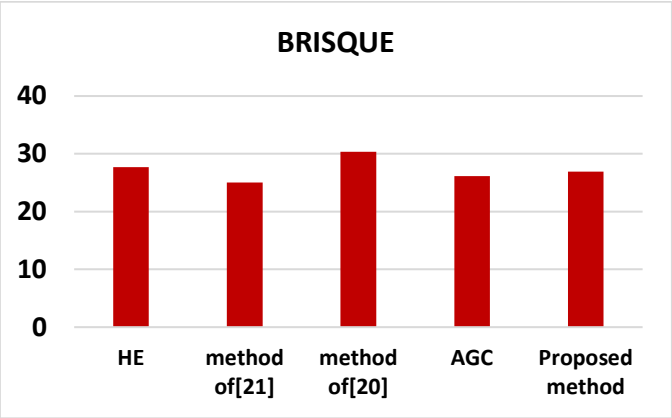


Figure 7. BRISQUE average for total images in a different method

Experimental results show the efficiency of our method in dark images is better than other methods. In contrast, the AGC method in the low-light image has better performance. Furthermore, as seen in the average case, the proposed method and AGC have similar performance in different metrics. In the SSIM metric, the proposed method has better performance than AGC; that is, the proposed method, better than AGC, preserves the structure of the image. Figure 8 shows some dark images of different enhancement methods compared to the proposed method. In this figure, (A) is the original image, (B) HE, (C) method of[21], (D) method of[20], (E) AGC, and (F) is the proposed method.

Figure 8. Results of different enhancement methods

Also, a comparison of the original images and the proposed method can be seen in Fig. 9.

(a) (b)

Figure 9. Original image (a) vs. Proposed method(b)

Figure 10 shows some low-light images of different enhancement methods compared to the proposed method. In this figure, (A) is the original image, (B) HE, (C) method of[21], (D) method of[20], (E) AGC, and (F) is related to the proposed method.

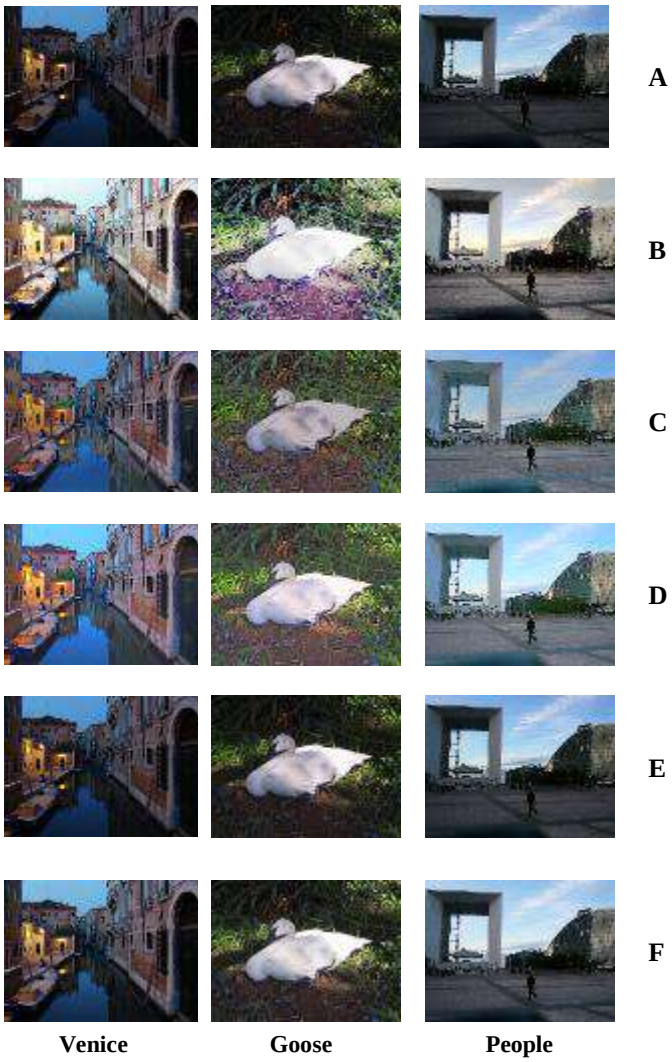


Figure 10. Results of different enhancement methods

Figure 11 shows a comparison of the original images and the proposed method for low-light images.

(a) (b)

Figure 11. Original image (a) vs. Proposed method(b)

V. CONCLUSION AND FUTURE WORK

Contrast enhancement is One of the most challenging issues in image processing. This issue, especially for low-light images with non-uniform illumination, is complex. In this paper, a method for improving contrast dark and low-light images based on the separation of red, green, and blue channels of an image and improving each channel, presented that while improving image contrast and noise reduction, the structural features of the image preserved to an acceptable level. In this research, the defined threshold is a fixed value that can be determined automatically based on the input image. Also, apply of this method for color images contrast enhancement using converted RGB to HSV space can be considered as a continuation of this research.

REFERENCES

[1] Alavi, Meysam, and Mehrdad Kargari. "A novel method for contrast enhancement of gray scale image based on shadowed sets." *2020 6th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS)*. IEEE, 2020.

[2] Maurya, Lalit, et al. "Contrast and brightness balance in image enhancement using Cuckoo Search-optimized image fusion." *Journal of King Saud University-Computer and Information Sciences* (2021).

[3] Cao, Gang, et al. "Contrast enhancement of brightness-distorted images by improved adaptive gamma correction." *Computers & Electrical Engineering* 66 (2018): 569-582.

[4] Gupta, Bhupendra, and Mayank Tiwari. "Minimum mean brightness error contrast enhancement of color images using adaptive gamma correction with color preserving framework." *Optik* 127.4 (2016): 1671-1676.

[5] Gupta, Bhupendra, and Tarun Kumar Agarwal. "New contrast enhancement approach for dark images with non-uniform illumination." *Computers & Electrical Engineering* 70 (2018): 616-630.

[6] Manju, R. A., G. Koshy, and Philomina Simon. "Improved method for enhancing dark images based on CLAHE and morphological reconstruction." *Procedia Computer Science* 165 (2019): 391-398.

[7] Srinivas, Kankanala, Ashish Kumar Bhandari, and Anurag Singh. "Low-contrast image enhancement using spatial contextual similarity histogram computation and color reconstruction." *Journal of the Franklin Institute* 357.18 (2020): 13941-13963.

[8] Akila, K., L. S. Jayashree, and A. Vasuki. "Mammographic image enhancement using indirect contrast enhancement techniques–a comparative study." *Procedia Computer Science* 47 (2015): 255-261.

[9] Kuran, Umut, and Emre Can Kuran. "Parameter selection for CLAHE using multi-objective cuckoo search algorithm for image contrast enhancement." *Intelligent Systems with Applications* 12 (2021): 200051.

[10] Kim, Taekyung, and Joonki Paik. "Adaptive contrast enhancement using gain-controllable clipped histogram equalization." *IEEE Transactions on Consumer Electronics* 54.4 (2008): 1803-1810.

[11] K. Zuiderveld, "Contrast limited adaptive histogram equalization," Graph. gems IV, pp. 474–485, 1994.

[12] Kim, Yeong-Taeg. "Contrast enhancement using brightness preserving bi-histogram equalization." *IEEE transactions on Consumer Electronics* 43.1 (1997): 1-8.

[13] Wang, Yu, Qian Chen, and Baeomin Zhang. "Image enhancement based on equal area dualistic sub-image histogram equalization method." *IEEE transactions on Consumer Electronics* 45.1 (1999): 68-75.

[14] Chen, Soong-Der, and Abd Rahman Ramli. "Contrast enhancement using recursive mean-separate histogram equalization for scalable brightness preservation." *IEEE Transactions on consumer Electronics* 49.4 (2003): 1301-1309.

[15] Sim, Kok Swee, Chih Ping Tso, and Yik Yee Tan. "Recursive sub-image histogram equalization applied to gray scale images." *Pattern Recognition Letters* 28.10 (2007): 1209-1221.

[16] Chen, Soong-Der, and Abd Rahman Ramli. "Minimum mean brightness error bi-histogram equalization in contrast enhancement." *IEEE transactions on Consumer Electronics* 49.4 (2003): 1310-1319.

[17] Ibrahim, Haidi, and Nicholas Sia Pik Kong. "Brightness preserving dynamic histogram equalization for image contrast enhancement." *IEEE Transactions on Consumer Electronics* 53.4 (2007): 1752-1758.

[18] Chiu, Yi-Sheng, Fan-Chieh Cheng, and Shih-Chia Huang. "Efficient contrast enhancement using adaptive gamma correction and cumulative intensity distribution." *2011 IEEE International Conference on Systems, Man, and Cybernetics*. IEEE, 2011.

[19] Huang, Shih-Chia, Fan-Chieh Cheng, and Yi-Sheng Chiu. "Efficient contrast enhancement using adaptive gamma correction with weighting distribution." *IEEE transactions on image processing* 22.3 (2012): 1032-1041.

[20] Li, Mading, et al. "Structure-revealing low-light image enhancement via robust retinex model." *IEEE Transactions on Image Processing* 27.6 (2018): 2828-2841.

[21] Fu, Xueyang, et al. "A fusion-based enhancing method for weakly illuminated images." *Signal Processing* 129 (2016): 82-96.

[22] Shi, Zhenghao, et al. "Nighttime low illumination image enhancement with single image using bright/dark channel prior." *EURASIP Journal on Image and Video Processing* 2018.1 (2018): 1-15.

[23] Li-Xin Wang, A Course in Fuzzy Systems and Control. Prentice Hall PTR, 1997.

[24] Wang, Zhou, et al. "Image quality assessment: from error visibility to structural similarity." *IEEE transactions on image processing* 13.4 (2004): 600-612.

[25] Bhandari, Ashish Kumar. "A logarithmic law based histogram modification scheme for naturalness image contrast enhancement." *Journal of Ambient Intelligence and Humanized Computing* 11.4 (2020): 1605-1627.

[26] Zhang, Lin, et al. "FSIM: A feature similarity index for image quality assessment." *IEEE transactions on Image Processing* 20.8 (2011): 2378-2386.

[27] Tan, Siu Fong, and Nor Ashidi Mat Isa. "Exposure based multi-histogram equalization contrast enhancement for non-uniform illumination images." *IEEE Access* 7 (2019): 70842-70861.

[28] Mittal, Anish, Rajiv Soundararajan, and Alan C. Bovik. "Making a "completely blind" image quality analyzer." *IEEE Signal processing letters* 20.3 (2012): 209-212.

[29] Mittal, Anish, Anush K. Moorthy, and Alan C. Bovik. "Blind/referenceless image spatial quality evaluator." *2011 conference record of the forty fifth asilomar conference on signals, systems and computers (ASILOMAR)*. IEEE, 2011.

[30] Wang, Wencheng, et al. "Adaptive image enhancement method for correcting low-illumination images." *Information Sciences* 496 (2019): 25-41.

[31] Wang, Wencheng, et al. "An experiment-based review of low-light image enhancement methods." *IEEE Access* 8 (2020): 87884-87917.

[32] Mittal, Anish, Anush Krishna Moorthy, and Alan Conrad Bovik. "No-reference image quality assessment in the spatial domain." *IEEE Transactions on image processing* 21.12 (2012): 4695-4708.

[33] Fu, Xueyang, et al. "A probabilistic method for image enhancement with simultaneous illumination and reflectance estimation." *IEEE Transactions on Image Processing* 24.12 (2015): 4965-4977.

[34] https://github.com/cs-chan/Exclusively-Dark-Image-Dataset.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

A New Edge and Pixel-Based Image Quality Assessment Metric for Colour and Depth Images

Seyed Muhammad Hossein Mousavi, S. Muhammad Hassan Mosavi

Independent Scientists, Tehran, Iran

***Abstract*— Measuring the quality of digital image is a complicated and importance task in image processing. Pixel and edge-based metrics are so crucial in dealing with a digital image. So, combination of edge and pixel features could handle not all but, almost all aspects of an image. Most recently using edge-based image quality metrics are popular, due to weakness of traditional image quality assessment metrics such as Peak Signal-to Noise Ratio. Also, most of the image quality metrics are belonged to color images, but recently new metrics for depth images are emerged. This paper proposes a new Full-Reference image quality assessment metric for color and depth images, which works based on edge and pixel features. Proposed method is a combination of improved Edge Based Image Quality Assessments and Peak Signal-to Noise Ratio methods. Proposed method is called Edge and Pixel-based Image Quality Assessment Metric (EPIQA). The system is validated using famous and benchmark performance metrics or quality measures such as Spearman Rank-Order Correlation Coefficient (SROCC), along with comparison with other similar methods on well-known related databases. Color databases have proper and diverse number of noises, but there is no proper depth noisy database, which it is decided to make one. Proposed method returned promising and satisfactory results in different tests.**

***Index Terms* —Image Quality Assessment (IQA) metric; Full-Reference (FR) metric; Edge and Pixel-based Image Quality Assessment (EPIQA); Depth Image Quality Metric**

I. INTRODUCTION

Fast technology progress, in the field of machine learning [51-52] and specially computer vision [50], increased the need for measuring techniques of images quality. The quality is the root characteristic of image, which assesses the amount of degradation and distortion such as blurriness, colour shift and all types of noises in a digital image. With having these values, it is possible to compare it with a perfect or reference image. It is applicable especially when for instance, a sensor wants to normalize the receiving image quality to an ideal version for human eyes. Furthermore, using an Image Quality Assessment (IQA) metric, helps to control and minimize the amount of receiving or transmitting distortions on images. The paper presents a new IQA metric to cover almost all-important aspects of a digital image and for different purposes. IQA methods play crucial roles in different applications of digital image processing like system performance validation, image improvement and restoration, noise removal, image compressions algorithms performance comparison, restoration, packet loss, and digital watermarking [1, 2, 3, 4, 5]. The paper is consisted of V sections, which are as follow. Section I, demonstrate fundamentals of research. Section II, pays to prior related researches on the subject. Section III, explains the proposed IQA metric in details. Section IV covers all experiments, comparisons and validation results. Also, section V includes, conclusion, dissuasion and future works.

A. Importance of the research

As it mentioned earlier, with increasing technological power of software and hardware in recent devices, it is needed to make more precise algorithms to deal with them. Digital image is not an exception in this area. Assessing the quality of digital image is important in three main fields. Initially, it can be employed for image quality monitoring for systems in quality control. For instance, a video acquisition system could get help from IQA metrics and adjust itself with the best receiving image possible. Second, IQA metric could be used in benchmark and famous systems, that are employed in image processing usages. Finally, this system can be mixed into an image processing machine to improve the algorithms' structure and the parameters.

A lot of traditional IQA metrics such as Mean squared error (MSE), Peak signal-to noise ratio (PSNR) and Structural similarity (SSIM) are completely use visual data inside an image matrix to calculate its quality. This information and data are too much as long as the Human Visual System (HVS) gets an image-based type data on its low-level information like the

edges, zero-crossing data, lines and corners [6, 7]. So, changes in low-level image features [53] are image degradations.

As it mentioned, most of the existing IQA metrics are made for specific condition, but this paper proposed a modified and combined version of an IQA metric which could overcome different image degradations. Also, proposed method could also be employed for depth or range images.

B. Human Visual System

As it is clear, Human Visual System (HVS) is an appropriate tool to assess the quality of an image, but it is not possible to be integrated with an electronic device and includes a lot of error due to changing from one person to another. Moreover, each person has different opinion and taste about visual quality, which makes it harder have a standard assessment based on human opinion. So, it is needed to have an electronic standard system based on HVS structure and even more [6, 7].

C. Image Quality Assessments (IQA) factor or metric

Any digital images could be polluted by a wide variety of noises and distortions in data transmission, data receiving, data computation, data compression, saving and this may result in quality prolusion and distortion. Image quality can refer to the level of accuracy that all imaging systems and sensors capture, process, store, compress, transmit and display [8]. As it mentioned, HVS is a fine tool for IQA, but it is not still a device to use. Automatic IQA metrics are made for these purposes. IQA metrics fall into two main categories. Subjective and objective IQA metrics.

C.1 Subjective

Image quality validation based on HVS is called subjective IQA. Some of the current subjective tests for IQA are done based on to Rec. ITUR BT.500-10 [9]. Plus, it is the best way to image quality assessment. The Mean Opinion Score or MOS is acquired by set of standard's results averaging. Actually, human eye observers rate the quality in terms of how bad the human thinks that the distortions are. Difference Mean Opinion Score (DMOS) is exactly as its name represents. Even if MOS can recognize Image Quality (IQ) better, it is time consuming process in real world usage [10, 11, 12]. Main point sizes of subjective validations are illustrated in Table I.

TABLE I. SUBJECTIVE ASSESSMENT OF IMAGES MAIN POINTS

MOS	1	2	3	4	5
Quality	Bad	Poor	Fair	Nice	Perfect
Deficiency	Very Bad	Bad	Slightly Bad	Sensible	Imperceptible

C.2 Objective

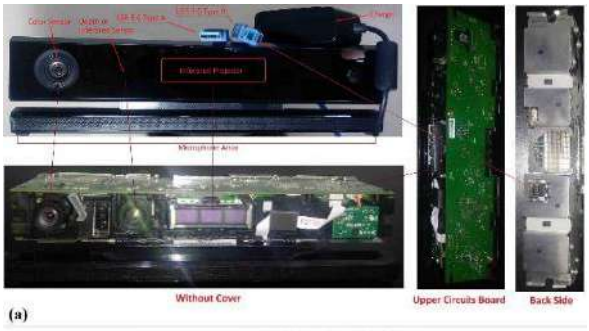
Makin an automatic IQA system which works as HVS is called objective IQA system. Many researches have been done during years to make a universal IQA system, but it was not so successful. These types of systems fall into three main categories. Full-Reference (FR), Non- Reference (NR) and Reduced-Reference (RR) [13]. This method is the most, accurate and famous one, which needs both reference and distorted images. Proposed method is belonged to this category [14]. It is called blind IQA. In this approach reference image is not exist. For instance, and in photography, mostly a NR algorithm is used to apprise the final user which a low-quality or high-quality image has been achieved. HVS could determine the quality in this condition, but it is so difficult to make an automatic system with same characteristics as human eyes [15]. This method is between FR and NR and just parts of the image is available. There are advantages like fewer data transmission, higher reliability, and wide usage range [16].

C.2.1 FR measures type

Full- Referenced (FR) IQA method could be classified into five classes of Pixel difference-based, Correlation-based, Edge-based, Context -based and Human Visual System-based systems as follow [17, 18]. As its name represents, every calculation is based on pixel by pixel. The mean square error (MSE), signal-to-noise ratio (SNR) and peak signal-to-noise ratio (PSNR) are in this class. Correlation is for comparing the difference of images. In IQA, pixels correlation is used as a scale of the quality. Here, original and the distorted images edges are found, then a displacement scale of edge parts is used to find the IQ in the entire digital image. In lieu of measuring pixels in both original and distorted data, pixel neighborhoods are compared by finding the probability to employee it for assessing the IQ.

D. Depth sensors and images

Depth sensors are made to calculate the distance between sensor and the object. Also, they can be employed to make 3-dimentional model of the object. With these abilities, they can be useful to increase the accuracy of the final recognition. Kinect is one of the most practical depth sensors to have. It is so much cheaper than other depth sensors and efficient. It can be used on Microsoft Xbox 360 (Kinect V.1) or Xbox one (Kinect V.2) consoles or be used as a developer device. Kinect 2.0 was released with Xbox One on November 22, 2013. Because of the lower price and high power to use, a lot of developers and researcher use it as a main depth device. It could record RGB and Depth video frames with 1920*1080 resolution for RGB images and 512*424 for Depth images on 30 fps. Moreover, it is also capable to of working between 0.8-5.0-meter ranges [19]. An RGB-D image is simply a combination of an RGB (color) image and its corresponding depth image. A depth image is an image channel in which each pixel relates to a distance between the image plane and the corresponding object in the RGB image. It is also termed as 2.5D or Range image [20, 46]. Fig 1 shows Kinect V.2 internal structure (a) and Kinect V.1 VS V.2 specifications (b). Fig 2 represents difference between color and depth images by a visual example (sample from proposed NDDB).



	Kinect V.1	Kinect V.2
Color	640*480 in 30 fps	920*1080 in 30 fps
Depth	320*240 in 30 fps	512*424 in 30 fps
Sensor	Structured Light	Time of Flight
Range	0.8 till 0.4 m	0.5 till 4.5 m
Angle of View	57/43 degree	70/60 degree
Microphone Array	yes	yes

Figure 1. Kinect V.2 internal structure (a) and Kinect V.1 VS V.2 specifications (b)

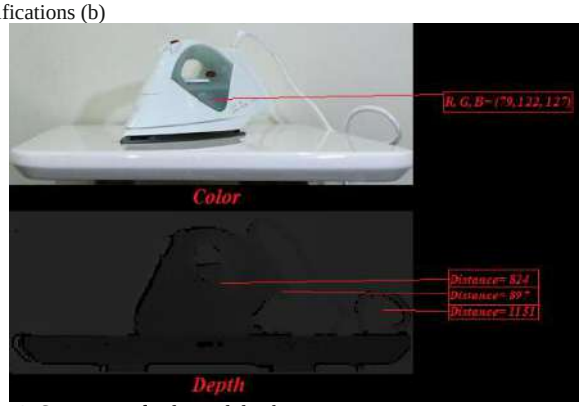


Figure 2. Structures of color and depth images

II. PRIOR RELATED RESEARCHES

In this section, some of the most famous and new IQA metrics (color and depth) are explained [47]. This metrics are mostly pixel and edge based.

A. PSNR

PSNR represents the unity level of signals. The Peak Signal to Noise Ratio or (PSNR) (1) matrix calculates the peak signal to noise ratio, in decibels unit, for images. The value is sometimes used as a IQ measurment for the original (input) and compressed (target) digital images. As higher the PSNR value, the better the compressing quality [12, 21].

$$\mathrm{PSNR} = 10\log_{10}\frac{L^2}{\mathrm{MSE}} \quad (1)$$

In that L is the range of value of pixel. It has unit of Dbwith limitation of 50. A good value is among 20 till 50.

B. MSE

The Mean Square Error or MSE (2) and PSNR are the two best error metrics in order to measure image compression quality. MSE shows the cumulative squared error for the compressed and the original image, which PSNR represents a measure for peak error. The lower value of MSE is better [21].

$$\mathrm{MSE} = \frac{1}{M \times N}\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}[X(i,j) - Y(i,j)]^2 \quad (2)$$

That, X and Y are arrays with size of M*N.

C. Structural Similarity (SSIM)

The calculation of SSIM is well documented in the references [22, 23], so here, paper gives only a brief review in order to be able to present the proposal performed and evaluated in this paper. Structural similarity can be achived by comparing pixel intensities' local template which is normalized for luminance and contrast factors. If f is the original image and g the distorted one, the SSIM (f, g) is calculated through the (3) equation.

$$\mathrm{SSIM} = \left(\frac{2\mu_f\mu_g + C_1}{\mu_f^2 + \mu_g^2 + C_1}\right)^{\alpha}\left(\frac{2\sigma_f\sigma_g + C_2}{\sigma_f^2 + \sigma_y^2 + C_2}\right)^{\beta}\left(\frac{\sigma_{fg} + C_3}{\sigma_f\sigma_g + C_3}\right)^{\gamma} \quad (3)$$

Being μ_i the mean intensity of the i image, σ_i their standard deviation and σ_{fg} is the covariance between the images.

D. Gradient Conduction Mean Square Error (GCMSE)

GCMSE is an edge aware metric based on MSE. Here, weighted sum of gradients (distance pixels) is considered [24]. GCMSE always brings better performance and result than MSE and SSIM. The steps are:

- Gradient directions are estimated in four aims using (4), and then value of G_p is found. Finally, The results are optimized by the k coefficient:

$$G = \frac{(I_2 - I_1)^2}{(I_2 - I_1)^2 + k^2} \quad (4)$$

The GCMSE is estimated based on (5):

$$\mathrm{GCMSE} = \frac{\sum_{x=1}^{m}\sum_{y=1}^{n}[(I_2(x,y) - I_1(x,y))G_p]^2}{C1 + \sum_{x=1}^{m}\sum_{y=1}^{n}G_p} \quad (5)$$

E. EBIQA

Actually, edge preservation is one of the most important aspects during the human visual assessment. Edge Based Image Quality Assessment (EBIQA) method is aim to act on the human understanding of the reciving features [25]. Steps are:

- Edge's locations are detect using Sobel's edge detection method in both images (original and destorted).

- A pixel window in size of 16 × 16 vectors are formed at each image based on (6) and (7), where I_1 is the reference image and I_2 is the test image.

$$I_1 = (O, AL, PL, N, VHO) \quad (6)$$
$$I_2 = (O, AL, PL, N, VHO) \quad (7)$$

In that 'O' presents orientation of edges in the digital image, that actually is the total number of edges. 'AL' is the length average of whole edges. 'PL' computes the total number of pixels which have similar intensity level. 'N' presents sum of all pixels, which shape all edges. 'VH' shows edges in either vertically or horizontally directions as sum of pixels of them. Finally, we estimate EBIQA by (8), which mean Euclidean distance of vectors is estimated. This paper going to improve this method and combine it with PSNR.

$$\mathrm{EBIQA} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\sqrt{(I_1 - I_2)^2} \quad (8)$$

F. NSER

NSER is a zero-crossings method based [26]. Steps are as follow:

- Using gaussian kernel in the images on different Standard Deviation (SD) gauge to detect all edges.
- Then procedure is done for both digital images. Common edge ratio number placed by first edge number is found by (9):

$$p_i = ||I_1 \cap I_2||/||I_1|| \quad (9)$$

- The result is normalized by log to optimize correlation (10):

$$\mathrm{NSER}(I_1, I_2) = -\sum_{i=1}^{N} \log 10(1 - p_i) \quad (10)$$

G. Measure of Enhancement or Enhancement Measure (EME)

This metric is based on the concepts of the Webers Low of the HVS. It helps to select the best parameters [27].

$$\mathrm{EME} = \max_{\Phi \in \{\Phi\}} \chi\big(\mathrm{EME}(\Phi)\big), \text{which } \max_{\Phi \in \{\Phi\}} = \chi\left(\frac{1}{k_1 k_2}\sum_{l=1}^{k_2}\sum_{k=1}^{k_1} 20 \log \frac{I^{\omega}_{max;k;l}}{I^{\omega}_{min;k;l}}\right) \quad (11)$$

Let an image x(n, m) be splitted into k_1k_2 blocks $w_{k,l}$ (i, j) of size $l_1 \times l_2$ and { Φ} be a provided class of orthogonal alter used for image improvement. Also $I^{\omega}_{max;k;l}$ and $I^{\omega}_{min;k;l}$ are respectively min and max of x (m, m) matrix inside the main block $w_{k,l.}$ The function χ is the sign one.

H. Full reference Feature Similarity Measure (FSIM)

Full reference feature similarity or (FSIM) index is to asses IQ [28]. FSIM can achive vital aspects of IQ by using two low-level characterstics. Those are phase and gradient magnitudes.

Le, Thanh-Ha, Seung-Won Jung, and Chee Sun Won proposed a new depth image quality metric which demands only a single pair of color and depth images in 2017. Their method closely estimates the depth quality metrics that use the ground truth depth or stereo color image pair [29].

Also, in 2013, Tsai, Chang-Ting, and Hsueh-Ming Hang, proposed a novel 3D IQ metric to assess the quality of stereo images that may contain artifacts introduced by the rendering process due to depth map errors [30].

Another valuable work is Tian, Shishun, et al research in 2018. They proposed a full-reference metric to assess the quality of Depth-Image-Based-Rendering (DIBR) synthesized views, as they believed 2D quality metrics may fail to evaluate the quality of the synthesized views [31]. Fig 3 represents the workflow of objective IQA metrics which proposed method is full reference version of it.

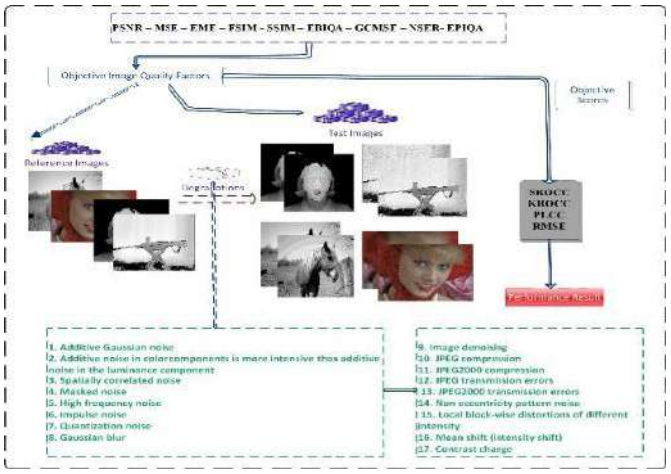


Figure 3. The workflow of objective Image Quality Assessment metrics

III. PROPOSED IQA FACTOR OR METRIC

As edges play important roles in detecting and distinguishing objects in HVS, and human eye has this ability or IQA power, it is decided to make an automatic IQA metric based on edge and pixel features, which works similar to human eyes. Also, for covering all details, a pixel-based approach is combined with the proposed approach. In other hands, using such technique on depth images is almost unique. Proposed metric is called Edge and Pixel-based Image Quality Assessment metric (EPIQA).

In order to do that, after image acquisition process from input, pre-processing stage starts. Pre-processing is consisting of low pass and high pass filtering. First, median filter (12) applies on both reference and distorted images, then un-sharp masking (13) effect takes place on them. This makes image, smooth from inside and sharp from outside, which help to decrease most type of noises effects. In this stage and at the same time, PSNR (1) value between distorted and reference images will be calculated. As it is obvious, PSNR value is in the range of 20-50, which for finale combination, this value normalizes between ranges of 0-1. Based on repeated experiments, using a custom filter for edge detection performs better than traditional edge detection algorithms in this research. So, edge detection stage, will be done using proposed filter presented in Table II to extract the edges. Next stage is belonging to dividing both distorted and reference images to 8*8 blocks. Using larger blocks, loses some of the information and smaller blocks make the system so slow. So, using 8*8 blocks is so rational. For each block in both images, two vectors of (14) and (15) are defined.

TABLE II. PROPOSED EDGE DETECTION FILTER [32]

-1.2	-0.8	-0.6		1.2	0.8	0.6
0	0	0		0	0	0
1.2	0.8	0.6		-1.2	-0.8	-0.6
1.2	0	-1.2		-1.2	0	1.2
0.8	0	-0.8		-0.8	0	0.8
0.6	0	-0.6		-0.6	0	0.6

- *Median filter*

Low-pass Median filter (12) is a nonlinear filtering technique, normally employed for noise removal in an image or a signal. The fundamental idea is to convolve on the image block by block, replacing each block with the median of neighboring entries [33].

$$y[m,n] = \text{median}\{x[i,j], (i,j) \in w \quad (12)$$

In which w shows a neighborhood, centered on location [m, n] in image.

- *Un-sharp masking*

A perfect sharpening technique which effects like a high-pass filter on digital image. Here opposite of un-sharp or not un-sharp filter which is masked (6) applies on image which means sharpening of edges. Parameters are amount, radius and threshold [34].

Un-sharp masking produce and image g(x, y) from an image f(x, y) via:

$$g(x,y) = f(x,y) - f_{smooth}(x,y) \quad (13)$$

Where $f_{smooth}(x,y)$ is smoother version of $f(x,y)$

$$O_{i,j}=(ED, ELA, GLR, NEP, EO \quad (14)$$

$$D_{i,j}=(ED, ELA, GLR, NEP, EO) \quad (15)$$

Which O and D are original and distorted images in i and j positions. Edge features are Edge Density (ED), Edge Length Average (ELA), Gray Level Region (GLR), Number of Edge Pixels (NEP) and Edge Orientation (EO).

Edge Density (ED): The number of edges in each 8*8 block of an image.

Edge Length Average (ELA): Average Length of each block. First, all edges' lengths are calculated, and a simple average makes it as a decimal number.

Gray Level Region (GLR): Number of regions with the same gray level in each block.

Number of Edge Pixels (NEP): Number of pixels for each edge in each block.

Edge Orientation (EO): Number of edges with vertical or horizontal orientation in each block.

For example, for $O_{i,j}$ and in Fig 4 (a) there are (6, 2, 3, 7, 5) values. Fig 4 (b) shows a sample from CSIQ database [35] in the edge detected form. In this Fig, right eye is shown in an 8*8 block. Such a block is used in each process of this paper.

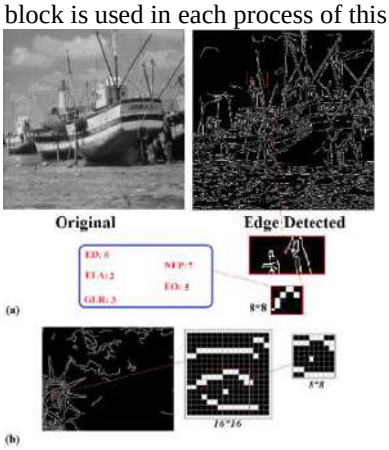


Figure 4. (a) Boat benchmark image (edge detection, final 8*8 sub block and extracting features), (b) a sample from CSIQ database (35) in the edge detected form. A sample of 8*8 block of reference image.

Now it is time to compute Euclidean distance $d_{i,j}$ between two corresponding blocks of original and distorted images, according to (16).

$$di,j = \left[\left(ED_{O_{i,j}} - ED_{D_{i,j}}\right)2 + \left(ELAO_{i,j} - ELAD_{i,j}\right)2 + \left(GLRO_{i,j} - GLRD_{i,j}\right)2 + \left(NEPO_{i,j} - NEPD_{i,j}\right)2 + \left(EOO_{i,j} - EOD_{i,j}\right)2\right]^{(\frac{1}{2})} \quad (16)$$

And average distance is computes by final EPIQA (17) in range of 0-1:

$$EPIQA: = 1 - \left(\frac{1}{M \times N \times MAX(d_{i,j})}\sum_{i=1}^{M}\sum_{j=1}^{N} d_{i,j}\right) \quad (17)$$

Final step is computing the average for EPIQA and PSNR values acquired from both images. Final value is a decimal number in range of 0 and 1. As it is clear, as the number is closer to 1, the quality is higher and vice versa. Fig 5 represents proposed metric's flowchart. Also, Fig 6 shows the proposed metric's steps in visual form using one of the samples from the proposed Noisy Depth database (NDDB) database. Proposed NDDB database is explained in the next section.

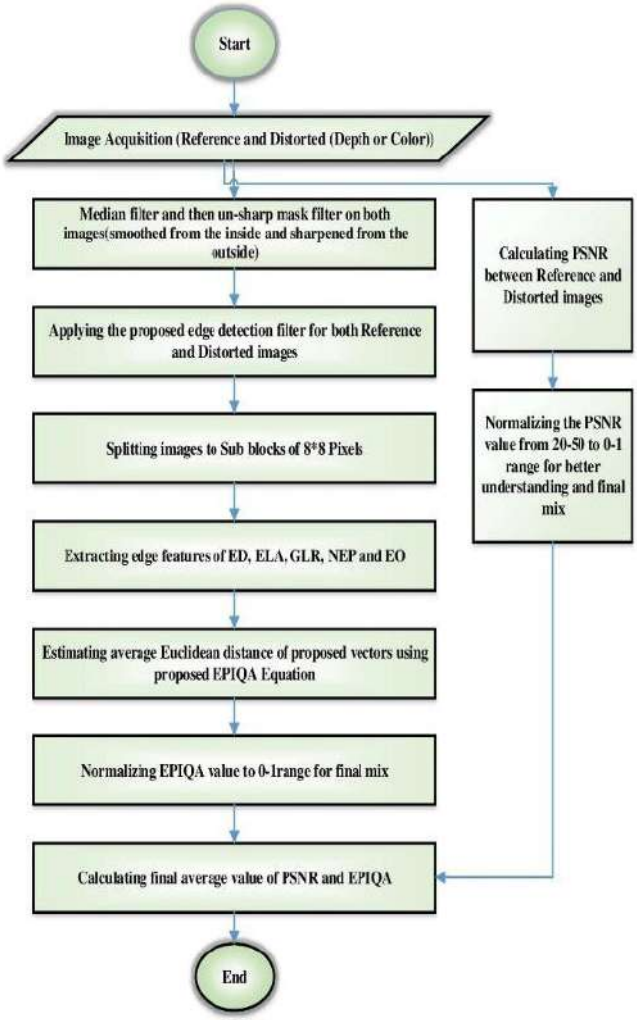


Figure 5. Proposed metric's flowchart

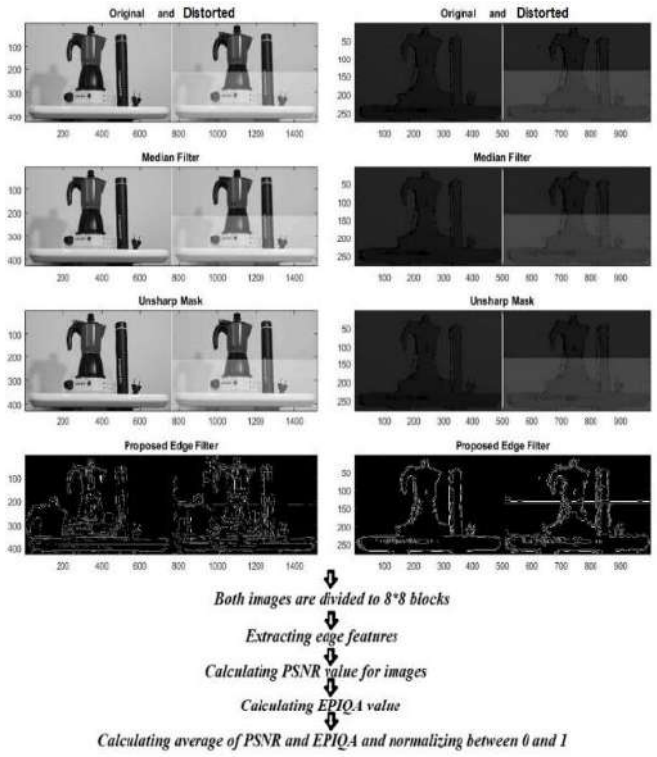


Figure 6. Proposed metric's steps in visual form (sample from proposed NDDB)

IV. VALIDATING EXPERIMENTAL RESULTS

System is evaluated using four databases. Two color and two depth databases, which one of them is proposed in this paper, that explained in follow. For depth databases, just depth data is used to show the efficiency of the system just on depth data. Furthermore, four validation metrics are used for final measurement, which are explained in follow. For this purpose, a system with following specifications is used: Intel(R) Core (TM) i7-4790K CPU @ 4.00GHz, 16GB RAM, NVIDIA GeForce GTX 1050, 233GB SSD in Matlab environment.

A. Databases

There are different color-based databases for IQA, which we use two of them in this research. Databases such as TID2008 [36], CSIQ [35], LIVE [37], IVC [38], MICT [39], A57 [1] are available, that A57 and TID2008 are used in this research. Also, for depth data sample from Eurecom Kinect [40] are used. In addition, a database which is consists of 30 depth images is made using Kinect version 2 sensor and proposed in this research. Proposed database is called Noisy Depth database (NDDB) and includes small home objects. This database is created due to the lack of desired noises in available databases.

For A57 database, pollutions are: FLT, JPG, JP2, DCQ, BLR and NOZ. For TID 2008 database, distortion types are 17 in 4 levels. Which makes 68 distorted images based on [36]. For depth images from depth databases, noises are added manually using seven noises of Gaussian [41, 42], Salt and pepper [41, 42], Poisson [43], Speckle [44], quantization [48], additive white gaussian [49] and block-wised [36] noises in just one level. These noises are added to Eurecom and proposed NDDB databases. Table III, presents the specifications of used databases. Additionally, some test samples from these databases are shown in Fig 7. Fig 8 shows recording environment for proposed NDDB database.

TABLE III. USED DATABASES' SPECIFICATIONS

	A57	TID2008	Eurecom	NDDB
Source images	3	25	52	30
Distorted images	54	1700	364	210
Distortion types	6	68 (17*4 levels)	7	7
Image type	Gray	Color	Depth	Depth
Size	512*512	512*384	256*256	512*424
Cite	[1]	[36]	[40]	-

Figure 7. Samples from used databases

Figure 8. Recording environment for NDDB database

B. Noises

Gaussian, salt and pepper, poison and speckle noises are explained in the paper and quantization, additive white gaussian and block-wised are referred to [48, 49] and [36] respectively and in order to save more space.

Gaussian noise is statistical noise with Probability Density Function (PDF) of the normal distribution. The PDF P of a Gaussian variable Z is given by (18):

$$P_g(Z) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(Z-\mu)^2}{2\sigma^2}} \quad (18)$$

In which Z shows the grey level, μ the mean value and σ represents standard deviation value.

Gaussian noise happens during data acquisition sensor noise caused by weak illumination and high environment temperature. Gaussian noise can be reduced using a spatial smoothing filter on the image [41, 42].

Salt-and-pepper or impulse noise (19) is a type of noise which rarely seen on digital images. This one can be caused by sharp and fast transmission in the image which shows itself as white and black pixels. Median filtering could fix this noise [41, 42].

$$P(Z) = \begin{cases} P_a & \text{for } z = a \\ P_b & \text{for } z = b \\ 0 & \text{otherwise} \end{cases} \quad (19)$$

Poisson noise (sometimes calls poisson noise) (20) is a type of noise which happens during electrical malfunctions, that can be defined by a Poisson formulation. In electronics shot noise originates from the discrete nature of electric charge [43].

$$Pr(N = k) = \frac{e^{-\lambda t}(\lambda t)^k}{k!} \quad (20)$$

Where λ is the number of photons per unit time interval.

Finally, Speckle noise (21) in conventional radar outcome from random sway in the return data of an object that is no bigger than a data image-processing element.

$$g(m,n) = f(m,n) * u(m,n) * \eta(m,n) \quad (21)$$

In which g (m, n) is distorted image, u (m, n) is multiplicative component and η (m, n) is extra component [44]. Effect of different noises on different images from used databases with different parameters is represented in Fig 9.

C. Validation metrics

For evaluating proposed EPIQA metric, four commonly performance metrics had been used. Used metrics are SROCC [25, 45], KROCC [25, 45], PLCC [25, 45] and RMSE [25, 45].

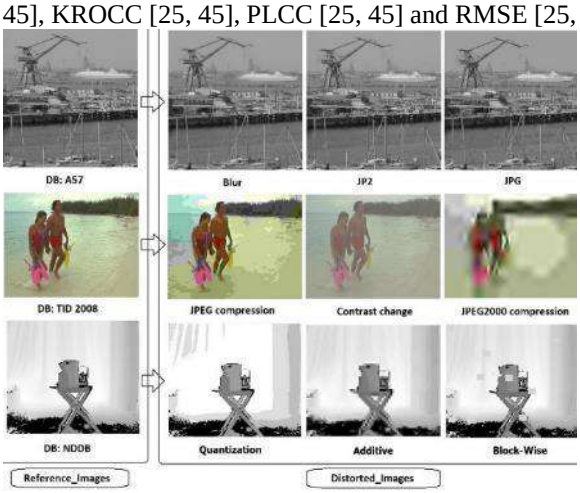


Figure 9. Effect of different noises on different images from used databases

C.1 Spearman rank-order correlation coefficient (SROCC)

SROCC shows the statistical association among the rankings of two variables. It evaluates how well the relationship between variables can be defined using a following function [25, 45].

$$SROCC = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)} \quad (22)$$

Where d_j is the difference between the i^{th} image's ranks in the subjective and objective evaluations.

C.2 Kendall rank-order correlation coefficient (KROCC)

This measure used to assess the serial dependence between two assessed quantities. Intuitively, the Kendall correlation between two variables will be high when observations have a homogeneous rank [25, 45].

$$KROCC = \frac{n_c - n_d}{0.5n(n-1)} \quad (23)$$

In which n_c and n_d are the number of compliant pairs and dissonant pairs in the data set.

C.4 Pearson Linear Correlation Coefficient (PLCC)

This one is the variables covariance which is divided by the product of their Standard Deviations (SD). In the following, it is supposed that s_j is the subjective score ofthe i^{th} image and x_j is the objective score ofthe i^{th} image. For the nonlinear regression analysis, first map x_i to q_i by the mapping function in (24) [25, 45].

$$q(x) = \beta_1\left(\frac{1}{2} - \frac{1}{1+e^{\beta_2(x-\beta_3)}}\right) + \beta_4 x \quad (24)$$

Where β_j are parameters which have to be fitted. Third one is PLCC between MOS and the objective scores after nonlinear regression.

$$PLCC = \frac{\sum_{i=1}^{n}(s_i - \bar{s})(q_i - \bar{q})}{\sqrt{\sum_{i=1}^{n}(s_i - \bar{s})^2}\sqrt{\sum_{i--1}^{n}(q_i - \bar{q})^2}} \quad (25)$$

C.5 Root-Mean-Square Error (RMSE)

Final metric is RMSE between MOS and the objective scores after nonlinear regression, which can be explained as (26) [25, 45].

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(s_i - q_i)^2} \quad (26)$$

D. Experimental results

System is validated using four common metrics of SROCC, KROCC, PLCC and RMSE on four databases of A57, TID2008, Eurecom and proposed NDDB, and compared with some of the traditional and new IQA metrics. All experiments are conducted under Matlab software environment. Tables IV to VII show acquired results on each database. As it is clear in all four databases, proposed EPIQA method could achieve better results in most cases. Also, the power of metrics could be categories as follow (from worst to best): SSIM, PSNR, [29], NSER, [30], FSIM, SC-IQA, EME, GCMSE, EBIQA and EPIQA. The neural linear regression with 70 hidden layers between proposed EPIQA and other methods and MOS is presented in Fig 10. These results are achieved for all databases and performance quality measures. Also, the results of performance metrics on all databases are represented and compared in Figs 11, 12, 13 and 14.

TABLE IV. COMPARISON RESULTS ON A57 DATABASE

	PSNR	SSIM	EBIQA	GCMSE	MSE	FSIM	EME	[29]	[30]	SC-IQA	NSER	EPIQA
SROCC	0.618	0.806	0.860	0.851	0.651	0.829	0.846	0.790	0.665	0.837	0.823	**0.874**
KROCC	0.530	0.605	**0.691**	0.620	0.563	0.628	0.632	0.636	0.577	0.673	0.609	0.666
PLCC	0.707	0.801	0.878	0.867	0.740	0.814	0.833	0.857	0.754	0.844	0.810	**0.901**
RMSE	0.173	0.146	0.119	0.135	0.145	0.139	0.133	0.137	0.149	0.104	0.120	**0.102**

TABLE V. COMPARISON RESULTS ON TID2008 DATABASE

	PSNR	SSIM	EBIQA	GCMSE	MSE	FSIM	EME	[29]	[30]	SC-IQA	NSER	EPIQA
SROCC	0.712	0.819	0.890	0.911	0.745	0.842	0.876	0.890	0.759	0.904	0.883	**0.916**
KROCC	0.690	0.688	0.763	0.727	0.703	0.711	0.728	0.742	0.717	0.770	0.725	**0.777**
PLCC	0.733	0.764	0.798	0.712	0.766	0.737	0.723	0.737	0.780	0.812	0.700	**0.823**
RMSE	0.127	0.126	0.114	0.120	0.125	0.129	0.138	0.112	0.139	**0.100**	0.125	0.106

TABLE VI. COMPARISON RESULTS ON EURECOM DATABASE

	PSNR	SSIM	EBIQA	GCMSE	MSE	FSIM	EME	[29]	[30]	SC-IQA	NSER	EPIQA
SROCC	0.699	0.709	**0.755**	0.720	0.712	0.712	0.730	0.734	0.716	0.749	0.727	0.742
KROCC	0.710	0.694	0.770	0.766	0.723	0.707	0.754	0.758	0.727	0.764	0.731	**0.775**
PLCC	0.728	0.717	0.749	0.727	0.711	0.730	0.743	0.737	0.725	0.763	0.730	**0.781**
RMSE	0.118	0.112	0.110	0.114	0.111	0.115	0.127	0.121	0.125	0.114	0.114	**0.106**

TABLE VII. COMPARISON RESULTS ON PROPOSED NDDB DATABASE

	PSNR	SSIM	EBIQA	GCMSE	MSE	FSIM	EME	[29]	[30]	SC-IQA	NSER	EPIQA
SROCC	0.751	0.740	0.766	**0.798**	0.744	0.743	0.759	0.763	0.758	0.780	0.756	0.783
KROCC	0.697	0.705	0.732	0.729	0.710	0.728	0.727	0.731	0.714	0.746	0.724	**0.750**
PLCC	0.730	0.704	0.735	0.731	0.723	0.727	0.737	0.738	0.747	0.740	0.734	**0.748**
RMSE	0.111	0.124	0.109	0.117	0.124	0.127	0.122	0.136	0.148	0.113	0.129	**0.101**

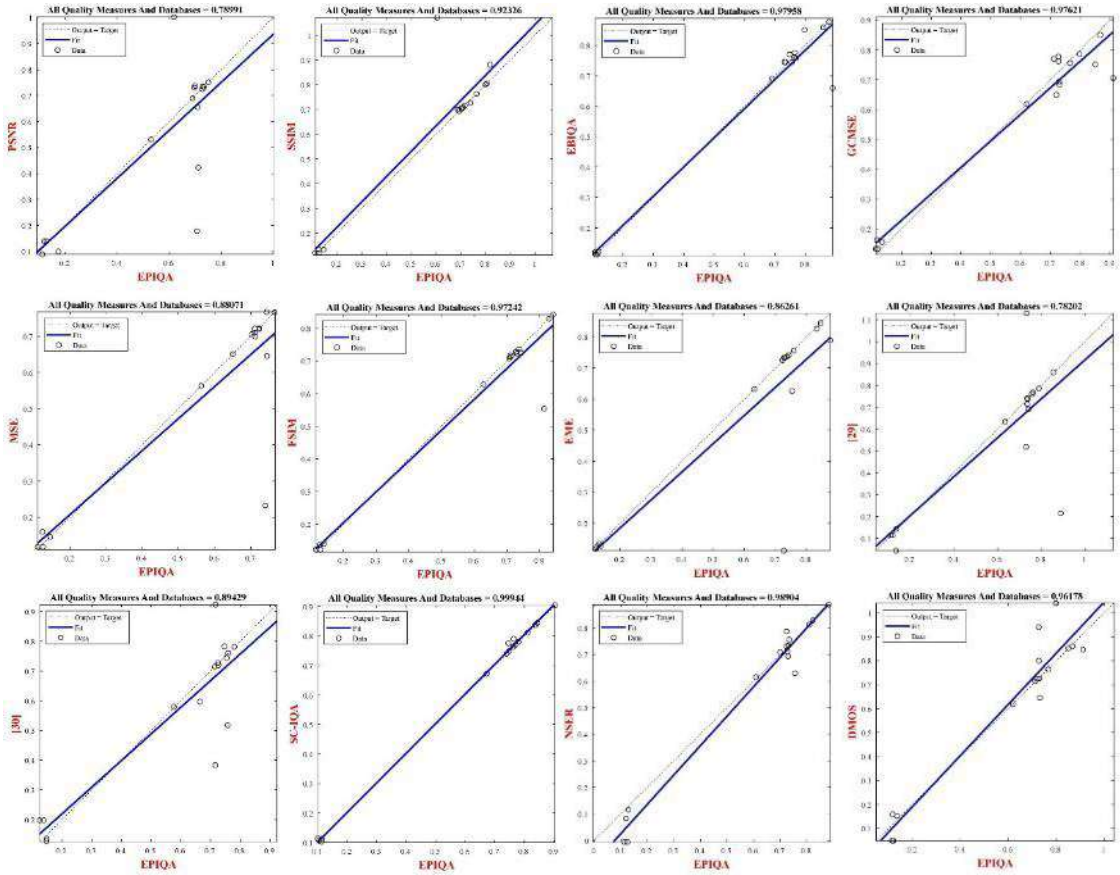


Figure 10. Neural linear regression for proposed method versus other methods (all databases and quality measures)

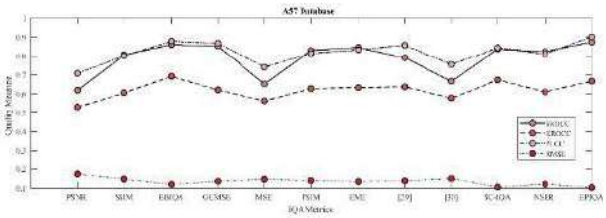


Figure 11. Performance result on A57 database

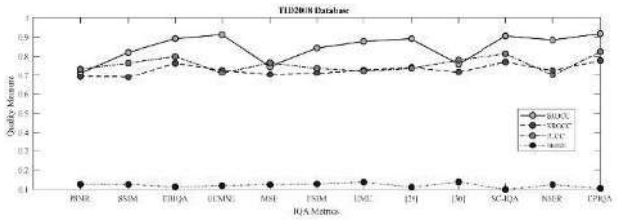


Figure 12. Performance result on TID2008 database

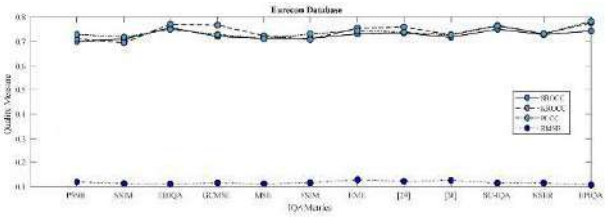


Figure 13. Performance result on Eurecom database

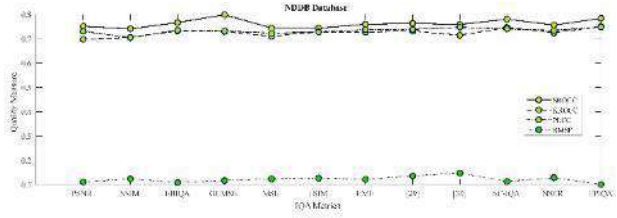


Figure 14. Performance result on NDDB database

V. CONCLUSION AND FUTURE WORKS

The paper proposed an edge and pixel based IQA metric for color and depth images. Proposed method was tested using commonly used performance metrics on well-known color and depth databases. Also due to lack of distorted depth database, a new database of NDDB is made for this purpose. Also, at comparison stage, proposed EPIQA metric is compared with some traditional and new metrics, which in both color and depth experiments, satisfactory results were achieved. Methods such as EME, SC-IQA, EBIQA and GCMSE had good performance after proposed method. Future work is consisting of adding new structure such as context-based metrics to exist version; and also making system compatible with NR and RR systems. It is suggested to make a new color and depth distorted database with more distortion types.

REFERENCES

[1] Chandler, Damon M., and Sheila S. Hemami. "VSNR: A wavelet-based visual signal-to-noise ratio for natural images." IEEE transactions on image processing 16.9 (2007): 2284-2298.

[2] Chandler, Damon M., Kenny H. Lim, and Sheila S. Hemami. "Effects of spatial correlations and global precedence on the visual fidelity of distorted images." Human Vision and Electronic Imaging XI. Vol. 6057. International Society for Optics and Photonics, 2006.

[3] Li, Xin. "Blind image quality assessment." IEEE International Conference on Image Proceedings. 2002. Vol. 1, 2002.

[4] Liu, Anmin, Weisi Lin, and Manish Narwaria. "Image quality assessment based on gradient similarity." IEEE Transactions on Image Processing 21.4 (2012): 1500-1512.

[5] Sheikh, Hamid R., Alan C. Bovik, and Gustavo De Veciana. "An information fidelity criterion for image quality assessment using natural scene statistics." IEEE Transactions on image processing 14.12 (2005): 2117-2128.

[6] Marr D (1982) "Vision: a computational investigation into the human rrepresentation and processing of visual information". W.H. Freeman and Company.

[7] Zhang, Min, et al. "Blind image quality assessment using the joint statistics of generalized local binary pattern." IEEE Signal Processing Letters 22.2 (2015): 207-210.

[8] Burningham, Norman, Zygmunt Pizlo, and Jan P. Allebach. "Image quality metrics." Encyclopedia of imaging science and technology (2002).

[9] BT, RECOMMENDATION ITU-R. "Methodology for the subjective assessment of the quality of television pictures." International Telecommunication Union (2002).

[10] Blasch, Erik, et al. "Image quality assessment for performance evaluation of image fusion." IEEE International Conference on Information Fusion, 2008.

[11] Navas, K. A., M. L. Aravind, and M. Sasikumar. "A novel quality measure for information hiding in images." 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. IEEE, 2008.

[12] Wang, Zhou, et al. "Image quality assessment: from error visibility to structural similarity." IEEE transactions on image processing 13.4 (2004): 600-612.

[13] Wang, Zhou, Eero P. Simoncelli, and Alan C. Bovik. "Multiscale structural similarity for image quality assessment." The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003. Vol. 2. Ieee, 2003.

[14] Sheikh, Hamid R., Muhammad F. Sabir, and Alan C. Bovik. "A statistical evaluation of recent full reference image quality assessment algorithms." IEEE Transactions on image processing 15.11 (2006): 3440-3451.

[15] Wang, Zhengyou, et al. "No-reference hybrid video quality assessment based on partial least squares regression." Multimedia Tools and Applications 74.23 (2015): 10277-10290.

[16] Yu, Ming, et al. "A method for reduced-reference color image quality assessment." 2009 2nd International Congress on Image and Signal Processing. IEEE, 2009.

[17] Sheikh, Hamid R., Muhammad F. Sabir, and Alan C. Bovik. "A statistical evaluation of recent full reference image qu ality assessment algorithms." IEEE Transactions on image processing 15.11 (2006): 3440-3451.

[18] Al-Najjar, Yusra AY, and D. Chen Soong. "Comparison of image quality assessment: PSNR, HVS, SSIM, UIQI." International Journal of Scientific & Engineering Research 3.8 (2012): 1.

[19] Sarbolandi, Hamed, Damien Lefloch, and Andreas Kolb. "Kinect range sensing: Structured-light versus time-of-flight kinect." Computer Vision and Image Understanding 139 (2015): 1-20.

[20] Ganguly, Suranjan, Debotosh Bhattacharjee, and Mita Nasipuri. "3D Image Acquisition and Analysis of Range Face Images for Registration and Recognition." Handbook of Research on Emerging Perspectives in Intelligent Pattern Recognition, Analysis, and Image Processing (2015): 185.

[21] ECE, CSE, and M. M. U. Mullana. "Image quality assessment techniques pn spatial domain." IJCST 2.3 (2011).

[22] Wang, Zhou, and Alan C. Bovik. "Modern image quality assessment." Synthesis Lectures on Image, Video, and Multimedia Processing 2.1 (2006): 1-156.

[23] Silvestre-Blanes, Javier, and Rubén Pérez-Lloréns. "SSIM and their dynamic range for image quality assessment." ELMAR, 2011 Proceedings. IEEE, 2011.

[24] López-Randulfe, Javier, et al. "A quantitative method for selecting denoising filters, based on a new edge-sensitive metric." 2017 IEEE International Conference on Industrial Technology (ICIT). IEEE, 2017.

[25] Attar, Abdolrahman, Asadollah Shahbahrami, and Reza Moradi Rad. "Image quality assessment using edge based features." Multimedia Tools and Applications 75.12 (2016): 7407-7422.

[26] Zhang, Min, Xuanqin Mou, and Lei Zhang. "Non-shift edge based ratio (NSER): An image quality assessment metric based on early vision features." IEEE Signal Processing Letters 18.5 (2011): 315-318.

[27] Agaian, Sos S., Karen Panetta, and Artyom M. Grigoryan. "A new measure of image enhancement." IASTED International Conference on Signal Processing & Communication. Citeseer, 2000.

[28] Zhang, Lin, et al. "Training quality-aware filters for no-reference image quality assessment." IEEE MultiMedia 21.4 (2014): 67-75.

[29] Le, Thanh-Ha, Seung-Won Jung, and Chee Sun Won. "A new depth image quality metric using a pair of color and depth images." Multimedia Tools and Applications 76.9 (2017): 11285-11303.

[30] Tsai, Chang-Ting, and Hsueh-Ming Hang. "Quality assessment of 3D synthesized views with depth map distortion." Visual Communications and Image Processing (VCIP), 2013. IEEE, 2013.

[31] Tian, Shishun, et al. "SC-IQA: Shift compensation based image quality assessment for DIBR-synthesized views." IEEE International Conference on Visual Communications and Image Processing. 2018.

[32] Mousavi, Seyed Muhammad Hossein and Marwa Kharazi. "An Edge Detection System for Polluted Images by Gaussian, Salt and pepper, Poisson and Speckle Noises.", 4th National Conference on Information Technology,Computer & TeleCommunication, At Iran-Torbat-e Heydarieh, 2017.

[33] Huang, T., G. J. T. G. Y. Yang, and G. Tang. "A fast two-dimensional median filtering algorithm." IEEE Transactions on Acoustics, Speech, and Signal Processing 27.1 (1979): 13-18.

[34] Polesel, Andrea, Giovanni Ramponi, and V. John Mathews. "Image enhancement via adaptive unsharp masking." IEEE transactions on image processing 9.3 (2000): 505-510.

[35] E. C. Larson and D. M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," Journal of Electronic Imaging, 19 (1), March 2010.

[36] N. Ponomarenko, V. Luki n, A. Zelensky, K. Egiazarian, M. Carli, F. Battisti, "TID2008 - A Database for Evaluation of Full-Reference Visual Quality Assessment Metrics", Advances of Modern Radioelectronics, Vol. 10, pp. 30-45, 2009.

[37] Sheikh HR, Sabir MF, Bovik AC (2006) "A statistical evaluation of recent full reference image quality assessment algorithms". IEEE Trans IP 15:3440–3451

[38] Subjective quality assessment IRCCyN/IVC database, https://qualinet.github.io/databases/image/irccynivc_image_quality_database/

[39] MICT Image Quality Evaluation Database, https://qualinet.github.io/databases/image/mict_image_quality_database/

[40] Rui Min, Neslihan Kose, Jean-Luc Dugelay, "KinectFaceDB: A Kinect Database for Face Recognition," IEEE Transactions on Systems, Man, and Cybernetics: Systems 44.11 (2014): 1534-1548.

[41] Gonzales, Rafael C., and Richard E. Woods. "Digital Image Processing", Addison & Wesley Publishing Company. Reading, MA (1992).

[42] Jain, Anil K. "Fundamentals of digital image processing". Englewood Cliffs, NJ: Prentice Hall, 1989.

[43] Hasinoff, Samuel W. "Photon, poisson noise." Computer Vision. Springer US, 2014. 608-610.

[44] Jaybhay, Jyoti, and Rajveer Shastri. "A study of speckle noise reduction Filters." signal & image processing: An international Journal (SIPIJ) 6.3 (2015): 71-80.

[45] Wang, Zhou, and Qiang Li. "Information content weighting for perceptual image quality assessment." IEEE Transactions on Image Processing 20.5 (2011): 1185-1198.

[46] Mousavi, Seyed Muhammad Hossein. "A New Way to Age Estimation for RGB-D Images, based on a New Face Detection and Extraction Method for Depth Images." International Journal of Image, Graphics and Signal Processing 10.11 (2018): 10.

[47] Mousavi, Seyed Muhammad Hossein, Vyacheslav Lyashenko, and Surya Prasath. "Analysis of a robust edge detection system in different color spaces using color and depth images." Компьютерная оптика 43.4 (2019).

[48] Robertson, Mark A., and Robert L. Stevenson. "DCT quantization noise in compressed images." IEEE Transactions on Circuits and Systems for Video Technology 15.1 (2005): 27-38.

[49] Damera-Venkata, Niranjan, et al. "Image quality assessment based on a degradation model." IEEE transactions on image processing 9.4 (2000): 636-650.

[50] Mousavi, Seyed Muhammad Hossein, and Vyacheslav Lyashenko. "Extracting old persian cuneiform font out of noisy images (handwritten or inscription)." 2017 10th Iranian Conference on Machine Vision and Image Processing (MVIP). IEEE, 2017.

[51] Mousavi, Seyed Muhammad Hossein, S. Younes MiriNezhad, and Atiye Mirmoini. "A new support vector finder method, based on triangular calculations and K-means clustering." 2017 9th International Conference on Information and Knowledge Technology (IKT). IEEE, 2017.

[52] Dezfoulian, Mir Hossein, et al. "Optimization of the Ho-Kashyap classification algorithm using appropriate learning samples." 2016 Eighth International Conference on Information and Knowledge Technology (IKT). IEEE, 2016.

[53] Mousavi, Seyed Muhammad Hossein, VB Surya Prasath, and Seyed Muhammad Hassan Mousavi. "Persian classical music instrument recognition (PCMIR) using a novel Persian music database." 2019 9th International Conference on Computer and Knowledge Engineering (ICCKE). IEEE, 2019.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

A Novel Approach to Optimizing the Initial Path of Mobile Robots in Static Environments

Mostafa Mirzaei[1], Aliakbar Niknafs[2]

[1] *Department of Computer Engineering, Shahid Bahonar University of Kerman, Kerman, Iran,* mirzaei@eng.uk.ac.ir
[2] *Department of Computer Engineering, Shahid Bahonar University of Kerman, Kerman, Iran,* niknafs@uk.ac.ir

***Abstract*—Path planning is one of the most important tasks for autonomous mobile robots. The purpose of path planning is to find a collision-free path from a starting position to a target position due to specific criteria such as distance. This paper proposes a modified path improvement algorithm based on the former and later points. In the initial path, some points are randomly placed on the line segments that have the necessary condition. New positions of points in the path are calculated by their adjacent points. This approach explores the most optimal path between different values of angle for the different number of segments. A method is proposed in which the initial path in the line segments that contain the turns adheres to the obstacles in the way. This operation makes a shorter path. Experiments are conducted on six maps and the results indicate that the proposed method achieves a considerable optimization in the final path.**

***Index Terms* —mobile robot, global path planning, autonomous navigation, particle swarm optimization, heuristics, motion planning**

I. INTRODUCTION

Nowadays automation is a popular research topic. Industrial development using intelligent mobile robots with advanced autonomous capabilities has been increased. Human tasks are replaced by robots to have advantages like speed, good accuracy, and more efficiency. For instance, using robots for harvesting [1], reconnaissance, spatial and terrestrial explorations [2] has significant usefulness. For other purposes like explosive ordnance disposal [3], and fighting against the Covid-19 pandemic [4], they are practical. Furthermore, employing mobile robots to carry heavy objects in working spaces would be useful.

Path planning is a basic concept to perform mentioned missions and one of the research domains in robotics that drew most researchers' attention after Lozano-Pérez and Wesley [5]. Path planning for a mobile robot can be considered as generating a collision-free path from a starting point to a target point and optimizing it concerning specific criteria.

Two categories that cover all approaches in robot path planning are off-line or global path planning and on-line or local path planning [6]. Global path planners must have recent and previous perceptive information of the spaces. Moreover, working space must be a known environment. Although the global path planner is insufficient to react in unknown maps or environments with dynamic obstacles, it is precious to generate an optimal path. Local path planners do not need past information from working spaces and usually obtain a path only over a fragment of global path based on information from on-board sensors. However, the local path planning algorithm is useless when the environment is cluttered or the target is located a long distance away, it works efficaciously in dynamic environments.

Methods in this area of research are mainly divided into heuristic and classical methods [7]. The most well-known classical methods are road map [8], potential field [9, 10], cell decomposition [11], and subgoal network [12]. The heuristic category contains fuzzy logic [13-15], neural network [16, 17], and the most famous nature-inspired methods such as particle swarm optimization [18-20] and genetic algorithm [21, 22]. Mentioned approaches have their own limitations, hence one independent algorithm cannot ideally find a solution for the robot path planning problem and the researchers attempt to find new integrated algorithms. A detailed review concerning robot navigation techniques can be found in [23-26].

Mobile robots have many different uses. Hence, various constraints and characteristics exist in their workspaces, consequently, there is not a unique definition for "optimal path" in all applications. In other words, there are several definitions for optimal path like a short distance, safety, smoothness, and least energy consumption that will be provided under the constraints such as same velocity and rotation angle.

This study presents a new method (MPI_FLP) to optimize the initial path of a mobile robot. The initial path of a mobile robot is obtained using SPS and Dijkstra algorithms. The proposed method which is inspired by the PI_FLP algorithm is applied to achieve an optimized path. Using MPI_FLP, the initial path in the line segments that contain the turns adheres to the obstacles in the way; thus, the optimal path concerning the length can be prepared.

The rest of the paper is organized as follows. In section II SPS and PI_FLP algorithms are reviewed. Then, the proposed method MPI_FLP is stated in section III. Results of simulation are given in section IV to show the effectiveness of the proposed method. Finally, section V concludes the paper.

II. SPS AND PI_FLP ALGORITHMS

Han and Seo [27] proposed a path planning method that has two steps. The first step is named surrounding point set (SPS) generates some points which encompass certain obstacles and the second step is path improvement by the former and latter points (PI_FLP), in which a new position for each point in the path is calculated by adjacent points. Adjacent points for a specific point in the path are the two points on either side. Generated path in the first step is called the initial path is further improved by the method that is proposed in the second step. In the second step, the initial path adheres to certain obstacles and is enhanced in terms of length and angle.

The first step of Han and Seo's approach is the algorithm to obtain the surrounding points set (SPS). This set of surrounding points is a series of points of a grid that encompasses certain obstacles known as critical obstacles. First, they used the method of connecting the starting point to the target point by a straight line to identify the critical obstacles out of all obstacles on the map. Any obstacle that is in contact with this straight line is called a critical obstacle. The number of critical obstacles is usually smaller than or equal to the total number of obstacles on the map. They apply the approach only to critical obstacles that is better than applying to all obstacles.

According to the slope of the straight line, a set of grid points is generated. Grid points can be placed in either SPS or tempset in each iteration to obtain the surrounding points set around the critical obstacles. When a grid point is considered as a reference point, it can be a member of SPS based on a specific criterion. Based on this criterion, if the reference point can be added to the SPS, the points of its adjacency are inserted into the tempset which is considered as a waiting list for the next reference points. On the other hand, if the reference point cannot satisfy the criterion, it cannot enter into the SPS, and its adjacent points cannot be considered as elements of the tempset. As mentioned earlier, the tempset is a queue and reference points will be respectively taken from the beginning of it.

The approach ends when the tempset is turned into a non-member set. This method is used for all critical obstacles, and when the surrounding points set is obtained for all these critical obstacles, the initial path from the starting point to the target point will be achieved by applying the Dijkstra algorithm.

The second step of Han and Seo's approach is the path improvement algorithm depending upon the former and later points (PI_FLP) that is a method for optimizing the initial path and has been inspired by the particle swarm optimization (PSO) algorithm. Han and Seo provided a method different from PSO concerning how to calculate the velocity vector. The new velocity vector of a certain point in the path is calculated by its former and later points. As mentioned earlier, Han and Seo's method in the second step optimizes the initial path. This will decrease the length of the path because the initial path tries to adhere to certain obstacles, and will increase the angles between two sequentially connected line segments; therefore, the smoothness of the path will increase.

In the second step of the method, a path is a set of points that are connected to each other from the starting point to the target point. i^{th} path in iteration t is expressed as (1):

$$X_{t,i} = \left\{ x_{t,i}(1), x_{t,i}(2), \ldots, x_{t,i}(N_i) \right\} \tag{1}$$

where $x_{t,i}(1) = S(x,y)$ is the starting point, $x_{t,i}(N_i) = T(x,y)$ is the target point, N_i is the number of points in path i, for all iterations and paths.

The velocity vector for a certain point for iteration t+1 is calculated by (2) and (3) as follows:

$$\left.\begin{aligned} v^f_{t,i}(p) &= x_{t,i}(p-1) - x_{t,i}(p) \\ v^l_{t,i}(p) &= x_{t,i}(p+1) - x_{t,i}(p) \end{aligned}\right\} \quad \forall p \in P \text{ and } p \neq S(x,y), T(x,y) \tag{2}$$

$$v_{t+1,i}(p) = \omega v_{t,i}(p) + r_1 v^f_{t,i}(p) + r_2 v^l_{t,i}(p) \tag{3}$$

where $x_{t,i}(p)$ is the position of certain point p in the path in particle i and iteration t, $x_{t,i}(p-1)$ is the former point and $x_{t,i}(p+1)$ is the later point of certain point p, ω is inertia weight that determines the amount of variation of the previous velocity vector, r_1, r_2 are random numbers with uniform distribution in the range [0, 1], P is the points' set of the initial path.

The new position of each point p in the path is calculated by (4):

$$x_{t+1,i}(p) = x_{t,i}(p) + v_{t+1,i}(p) \quad \forall p \in P \text{ and } p \neq S(x,y), T(x,y) \tag{4}$$

where $S(x,y)$ is the start point and $T(x,y)$ is the target point in the path.

There is a possibility of obstacles' collision for some line segments of the path. So, after updating the new position for all points in a path, if a collision occurs, a collision removing process will be implemented by (5) as follows:

$$\begin{aligned} x_{t+1,i}(AP) = {} & x_{t,i}(AP) \\ & + \lambda \left(x_{t+1,i}(AP) - x_{t,i}(AP) \right) \\ & \text{where } 0 < \lambda < 1 \end{aligned} \tag{5}$$

where AP is a set of points that their line segments collide with obstacles.

This method could remove the collision and also applies the least change from the current position. Fig. 1 illustrates the operation of PI_FLP.

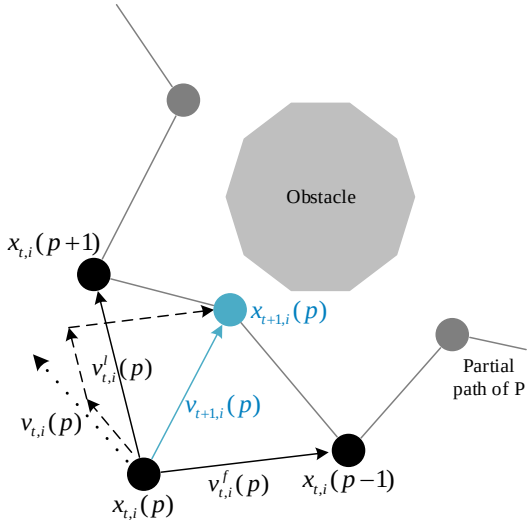


Fig. 1. Updating new position

The cost function in Han and Seo's approach is the path length and is based on the sum of distances. So, it is calculated by summing all $N_i - 1$ line segments in the path. Therefore, the best path in iteration t is the i^{th} particle if it has the smallest value of (6).

$$Distance_{t,i} = \sum_{p=1}^{N_i-1} \left\| x_{t,i}(p) - x_{t,i}(p+1) \right\| \qquad (6)$$

III. Proposed Method

The proposed method, modified path improvement algorithm based on the former and later points (MPI_FLP), is inspired by the second step of Han and Seo's approach. In the proposed method, a path is expressed as (1), and for calculating the velocity vector and new position of a point in the path, we use (2), (3), and (4). Because of collision removing with (5) applies the least change from the current position [27], it is also used in the proposed method to remove the collision.

In the MPI_FLP algorithm, the initial path is copied into several particles. Each point in the path is a feature in the particle and consequently has its velocity vector. After doing a process such as placing the initial path in particles, calculating the points' velocity vector, repositioning of points, and removing the collisions, there would be a condition in some points which are located in the turning parts of the path. Therefore, Optimization using the former and later points is not possible due to the lack of sufficient points in the turning parts. For instance, the new position of point p is calculated by (2), (3), and (4). Fig. 2a indicates this operation.

As it is clear from Fig. 2a, there is an obstacle collision at the new position of point p. When the collision is removed using (5), a new position of point p and a new path will be achieved. Fig. 2b shows the correction of collision with the obstacle. As it is illustrated in Fig. 2b, point p in its new position due to its obstacle collision is known as point *AP*. When the collision is removed, the new position of point *AP* constructs the line segments 1 and 2. This part of the path is not able to achieve further improvement which consequently cannot allow the robot to get closer to the obstacles in the next iterations. In other words, line segment 2 collides with the obstacle in the next iterations and according to (5), point *AP* returns to its approximate previous place. Thus more iteration to obtain better optimization is useless and calculations will be wasted.

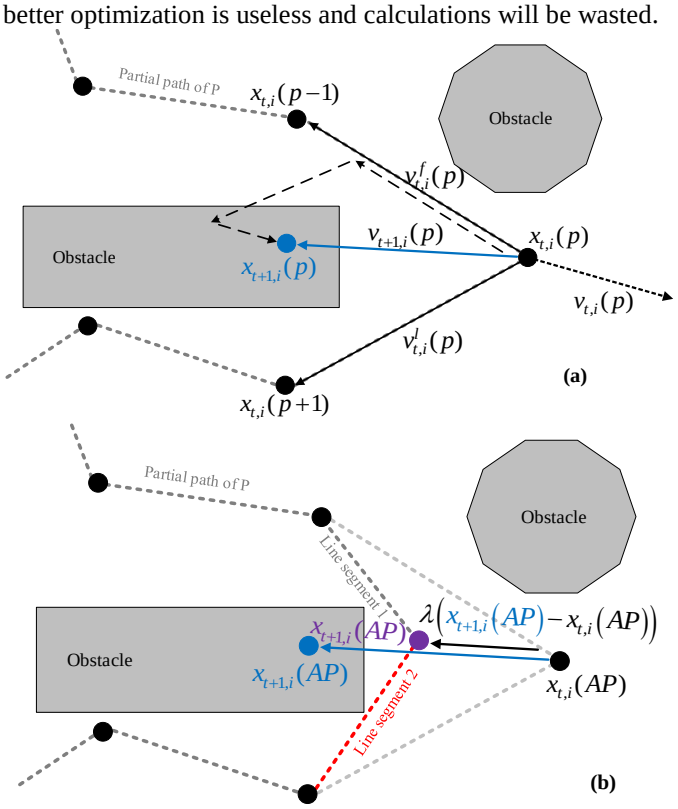


Fig. 2. Removing the collision

Each particle is an answer in the population, containing the initial path from the starting point to the target point. Each line segment that contains a turn is segmented. Firstly, the angle θ between all two line segments is calculated. The line segments will be segmented if the calculated angle θ is equal to or smaller than β. In other words, some random points will be considered on all two line segments that have the mentioned condition for angle θ. Hence, the number of path points in the line segment which has a turn with the specified condition increases. The method uses uniform distribution to generate random points. Fig. 3 shows the segmentation of a line segment in the partial path of P.

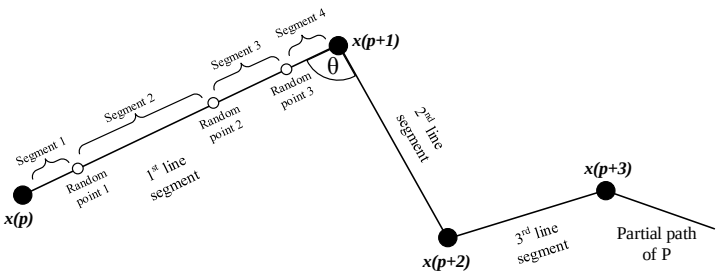


Fig. 3. The segmentation of a line segment in the partial path of P

This operation is considered for all population particles. So, the line segments are probably not the same in extra points. For example, the angle θ between two line segments in the initial path is calculated for three partitions. If angle θ for every two line segments is equal to or smaller than β, two points on these line segments are randomly placed in all particles. Thus the number of points increases to seven for these two line segments in all particles. Fig. 4 indicates this operation for two particles. It is obvious that the number of random points is one unit less than the number of segments in each line segment.

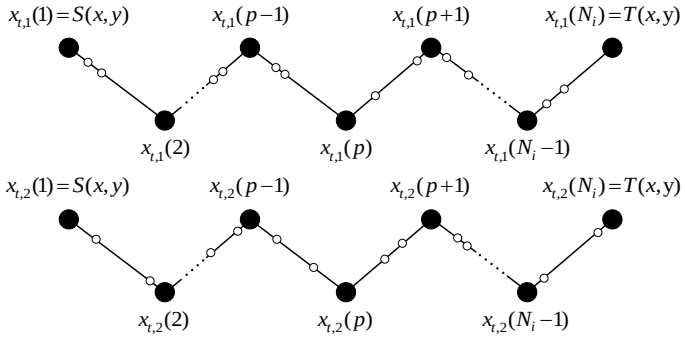


Fig. 4. Initialization of two particles

In the initialization step, the segmentation is performed after the angle θ is calculated. To calculate the angle θ between two line segments in the path, (7) [27] is used.

$$\text{Angle} = \cos^{-1}\left(\frac{v_{t,i}^{f}(p).v_{t,i}^{l}(p)}{\left\|v_{t,i}^{f}(p)\right\|\left\|v_{t,i}^{l}(p)\right\|}\right) \quad (7)$$

The proposed approach explores the most optimal path which is the shortest path between different angles for the different number of segments. When segmentation based on the value of θ for all line segments of the initial path is finished, the new position for each point in the path is calculated using (2), (3), and (4). Also, if a collision occurs at the new position, it is removed by (5). After calculating the new positions and removing the collisions for all points in all particles, the value of the cost function for each particle is calculated using (6) and the best path in the current iteration will be stored.

The continuation of iterations stops when there is no improvement in the best path in several consecutive iterations. Then, the shortest path achieves as the best-stored path for the current value of angle θ. Note that the value of angle θ decreases to the amount of δ. Then, particles are initialized again same as stated before based on the new value of θ. This operation continues until the angle θ reaches the considered smallest value.

Finally, the best-stored paths with different angels are searched to obtain the best path for the current value of segmentation. This operation is performed for each value of segmentation.

As an effective method to optimize the initial path, we adjust the inertia weight using (8). Hence, in the proposed method the inertia weight varies by the number of iterations.

$$\omega = \omega_{\max} - \frac{(\omega_{\max} - \omega_{\min}) * \mathrm{z}}{\max ite} \quad (8)$$

where ω_{max} is the considered maximum value for inertia weight in the algorithm, ω_{min} is the considered minimum value for inertia weight, z is the current iteration and $\max ite$ is the maximum number of iteration in the algorithm.

IV. SIMULATION

To evaluate the proposed optimizer, we simulated and executed it on a system equipped with Windows 8, a central processing unit of model Intel® Core™ i3-3110M with a frequency of 2.40GHz and main memory with the amount of 4GB.

Six maps are chosen for simulations which are different in terms of the number of obstacles, the shape of obstacles, and the width of roads (http://imr.ciirc.cvut.cz/planning/maps.xml). Start and target points which are randomly located in the free space in [27] are also used in the relevant maps here. The number of obstacles on the maps, start, and target points are depicted in Table I. Experiments are performed for each map to optimize the initial path from the starting point to the target point.

As mentioned earlier, the cost function in the proposed approach is the path length and the best path has the shortest distance from the starting point to the target point. All experiments are evaluated based on the path length. Path smoothness is also calculated based on the average angle for all points except the first and the last one in the path and is presented in Table II.

Table II describes the best path for the different number of segments and the output of the proposed method is specified with a grey row for each map. The output of the proposed method for map (1) is a path that is segmented by three and has a length that decreased from 2710.83 to 2665.2771 compared to the initial path. The best angle to segment is 162 degrees and the number of points on the optimal path is increased from 16 to 40 relative to the initial path. The average smoothness for the optimized path is 166.7987 that is calculated for 38 turns.

Table I. Details of the maps, start and target points

Map No.	Map name	Start		Target		No. of obstacles
		x	y	x	y	
1	Corridor_wavy	253.64	403.56	1895.31	1206.97	2
2	Bugtrap1	454.4	359.98	1003.6	608.77	1
3	Complex	615.78	886.62	1909.13	1425.31	6
4	Potholes	307.29	1602.93	1906.04	500.98	23
5	Rockpile	303.9	1609.04	1808.15	500.84	25
6	Var_density	1716.31	1501.95	607.37	406.88	42

For map (6), the output of the algorithm is the path that is segmented by three. In this path, the best angle for segmenting required line segments is 180 degrees and the number of points in this optimal path has increased from 7 to 19 compared to the initial path. The length of this path has decreased from 1632.47

to 1576.1057 relative to the initial path and the average smoothness for 17 turns is 173.7284. The results for other maps are presented in Table II.

The output of the proposed algorithm for all maps is illustrated in Fig. 5. As shown in this figure, the final path is specified with a solid red line.

As can be seen from the results of the proposed method, an increase of the points in the line segments which have turns would affect on proximity of these line segments to the obstacles in the way; thus, better optimization in turns would shorten the final path.

Table II. Simulation results of MPI_FLP. The output of MPI_FLP is specified with the grey row.

Map No.	Number of points in the initial path	Length of the initial path	Average of smoothness in the initial path	Number of segments	Best angle	Iteration	Number of points in the optimized path	Length of optimized path	Average of smoothness in the optimized path
1	16	2710.83	151.57	1	180	35	16	2694.336	151.8384
				2	180	24	31	2676.9494	165.4433
				3	162	37	40	2665.2771	166.7987
				4	162	56	52	2666.2369	170.1291
				5	162	39	64	2671.7209	163.7181
2	7	764.77	155.32	1	180	42	7	757.9624	156.1291
				2	126	16	12	751.6468	160.1042
				3	108	25	13	749.4769	168.6234
				4	108	19	16	750.4337	167.6156
				5	108	32	19	751.4756	168.4063
3	6	2954.46	112.74	1	180	38	6	2812.8609	126.414
				2	144	38	11	2799.3338	154.0781
				3	180	33	16	2792.9453	161.4654
				4	126	40	21	2792.8775	164.158
				5	144	63	26	2791.8384	168.1822
4	5	1993.26	154.60	1	180	39	5	1951.3861	173.3512
				2	144	38	9	1948.3245	175.5865
				3	144	55	13	1948.2241	177.3524
				4	54	62	5	1949.7379	173.6141
				5	90	61	5	1949.9799	173.4298
5	13	2067.55	155.24	1	180	43	13	2048.1912	160.083
				2	180	20	25	2030.7037	168.2217
				3	180	44	37	2036.2401	170.2971
				4	162	35	43	2034.273	169.6085
				5	162	17	53	2046.7724	161.4694
6	7	1632.47	142.23	1	180	114	7	1583.3954	153.2508
				2	162	31	13	1577.869	170.3296
				3	180	82	19	1576.1057	173.7284
				4	144	45	22	1578.2252	172.3283
				5	126	62	27	1578.3363	170.2929

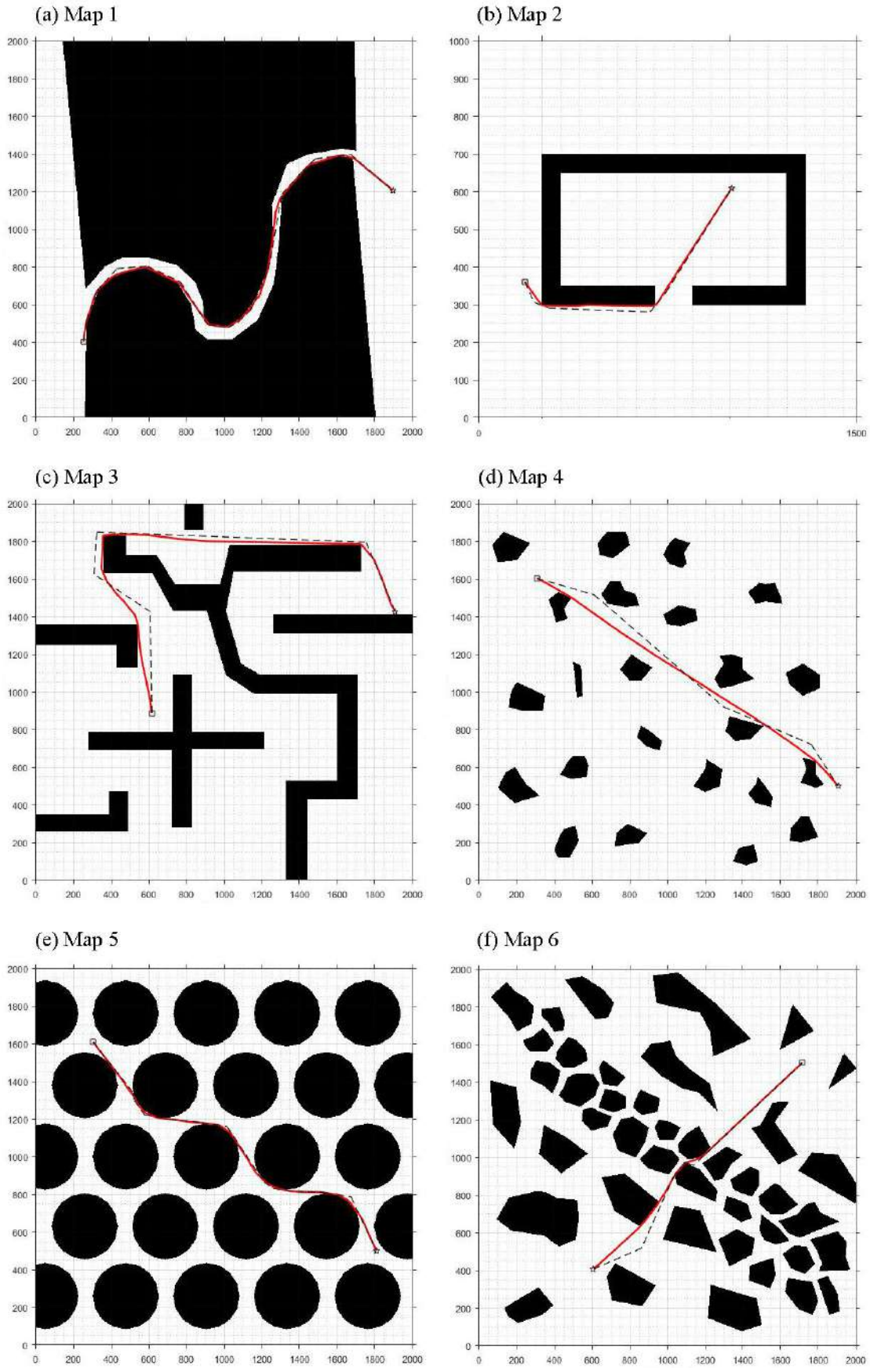


Fig. 5. Output of MPI_FLP. Start point, target point, initial path and output of MPI_FLP algorithm are illustrated by black square, black star, dashed black line and solid red line respectively.

V. CONCLUSION

This study proposes the MPI_FLP method for shortening the initial path of mobile robots. The PI_FLP algorithm is applied to the method. Each point in the initial path is considered as a feature in particles, and the velocity vector is calculated for all points. Moreover, the new position is updated based on the former and later points.

The proposed method dramatically optimizes the initial path based on the length. This algorithm tries to achieve the most optimal path in line segments that contain turns. Such optimization occurs because line segments that contain turns try to adhere to the obstacles in the way. This method can optimize the path in very narrow spaces and it is not important how small the free space is. The method can be effective to reduce the role of less random property. The random numbers considered in this method show the least role in calculating velocity vectors. The boundaries of every obstacle can be expanded to collision avoidance. The size of the robot plus a safe distance can be considered as the amount of expansion.

Future studies are planned to determine the number of segments and appropriate value for δ which can enhance the efficiency of the proposed method.

REFERENCES

[1] C. W. Bac, T. Roorda, R. Reshef, S. Berman, J. Hemming, and E. J. van Henten, "Analysis of a motion planning problem for sweet-pepper harvesting in a dense obstacle environment," *Biosystems Engineering*, vol. 146, pp. 85-97, 2016.

[2] J. Vallvé and J. Andrade-Cetto, "Potential information fields for mobile robot exploration," *Robotics and Autonomous Systems*, vol. 69, pp. 68-79, 2015.

[3] W. Deng, H. Zhang, Y. Li, and F. Gao, "Research on target recognition and path planning for EOD robot," *Int. J. Computer Applications in Technology*, vol. 57, no. 4, pp. 325-333, 2018.

[4] M. Cardona, F. Cortez, A. Palacios, and K. Cerros, "Mobile robots application against Covid-19 pandemic," in *2020 IEEE ANDESCON*, 2020, pp. 1-5: IEEE.

[5] T. Lozano-Pérez and M. A. Wesley, "An algorithm for planning collision-free paths among polyhedral obstacles," *Communications of the ACM*, vol. 22, no. 10, pp. 560-570, 1979.

[6] P. Raja and S. Pugazhenthi, "Optimal path planning of mobile robots: A review," *International Journal of Physical Sciences*, vol. 7, no. 9, pp. 1314-1320, 2012.

[7] E. Masehian and D. Sedighizadeh, "Classic and heuristic approaches in robot motion planning-a chronological review," *World Academy of Science, Engineering and Technology*, vol. 23, no. 5, pp. 101-106, 2007.

[8] A. N. Nazif, A. Davoodi, and P. Pasquier, "Multi-agent area coverage using a single query roadmap: A Swarm Intelligence Approach," in *Advances in practical multi-agent systems*, vol. 325: Springer, 2010, pp. 95-112.

[9] H. Yunqiang, K. Wende, C. Lin, and L. Xiaokun, "Research on multi-objective path planning of a robot based on artificial potential field method," *Int. J. Wireless and Mobile Computing*, vol. 15, no. 4, pp. 335-341, 2018.

[10] N. H. Singh, S. S. Devi, and K. Thongam, "Modified artificial potential field approaches for mobile robot navigation in unknown environments," in *Soft Computing for Problem Solving*: Springer, 2020, pp. 319-328.

[11] N. Ghita and M. Kloetzer, "Trajectory planning for a car-like robot by environment abstraction," *Robotics and Autonomous Systems*, vol. 60, no. 4, pp. 609-619, 2012.

[12] N. N. Singh, A. Chatterjee, A. Chatterjee, and A. Rakshit, "A two-layered subgoal based mobile robot navigation algorithm with vision system and IR sensors," *Measurement*, vol. 44, no. 4, pp. 620-641, 2011.

[13] C. Rekik, M. Jallouli, and N. Derbel, "Optimal trajectory of a mobile robot using hierarchical fuzzy logic controller," *Int. J. Computer Applications in Technology*, vol. 53, no. 4, pp. 348-357, 2016.

[14] M. s. Boujelben, C. Rekik, and N. Derbel, "A reactive approach for mobile robot navigation in static and dynamic environment using fuzzy logic control," *Int. J. Modelling, Identification and Control*, vol. 27, no. 4, pp. 293-302, 2017.

[15] G. Na, L. Caihong, W. Di, S. Yong, G. Tengteng, and L. Guoming, "Path Planning of Mobile Robot Based on Improved Fuzzy Control," in *2020 39th Chinese Control Conference (CCC)*, 2020, pp. 3671-3676: IEEE.

[16] S. Parasuraman and S. C. Yun, "Mobile robot navigation: neural Q–learning," *Int. J. Computer Applications in Technology*, vol. 44, no. 4, pp. 303-311, 2012.

[17] N. Guo, C. Li, D. Wang, Y. Song, G. Liu, and T. Gao, "Local Path Planning of Mobile Robot Based on Long Short-Term Memory Neural Network," *Automatic Control and Computer Sciences*, vol. 55, no. 1, pp. 53-65, 2021.

[18] X. Xu, Y. Guan, B. Hong, W. Ke, and S. Piao, "A humanoid robot path planning method based on virtual force-directed particle swarm optimisation," *Int. J. Wireless and Mobile Computing*, vol. 9, no. 4, pp. 325-331, 2015.

[19] B. Song, Z. Wang, and L. Zou, "An improved PSO algorithm for smooth path planning of mobile robots using continuous high-degree Bezier curve," *Applied Soft Computing*, vol. 100, p. 106960, 2021.

[20] J. Zhao, X. Sheng, and J. Shi, "Path Planning of Mobile Robot Based on Simulated Annealing Particle Swarm Optimization Algorithm," in *International Conference on Green Communications and Networking*, 2020, pp. 131-139: Springer.

[21] M. Davoodi, F. Panahi, A. Mohades, and S. N. Hashemi, "Clear and smooth path planning," *Applied Soft Computing*, vol. 32, pp. 568-579, 2015.

[22] P. G. Luan and N. T. Thinh, "Hybrid genetic algorithm based smooth global-path planning for a mobile robot," *Mechanics Based Design of Structures and Machines*, pp. 1-17, 2021.

[23] F. Gul, W. Rahiman, and S. S. Nazli Alhady, "A comprehensive study for robot navigation techniques," *Cogent Engineering*, vol. 6, no. 1, p. 1632046, 2019.

[24] K. Karur, N. Sharma, C. Dharmatti, and J. E. Siegel, "A Survey of Path Planning Algorithms for Mobile Robots," *Vehicles*, vol. 3, no. 3, pp. 448-468, 2021.

[25] B. B. K. Ayawli, R. Chellali, A. Y. Appiah, and F. Kyeremeh, "An overview of nature-inspired, conventional, and hybrid methods of autonomous vehicle path planning," *Journal of Advanced Transportation*, vol. 2018, 2018.

[26] B. Patle, A. Pandey, D. Parhi, and A. Jagadeesh, "A review: On path planning strategies for navigation of mobile robot," *Defence Technology*, vol. 15, no. 4, pp. 582-606, 2019.

[27] J. Han and Y. Seo, "Mobile robot path planning with surrounding point set and path improvement," *Applied Soft Computing*, vol. 57, pp. 35-47, 2017.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

A Novel Control Strategy Based on Fuzzy Logic in Islanded Microgrid

Farzad Rastegar

Department of Electrical Engineering, Sari Branch, Islamic Azad University, Sari, Iran, Email: farzadrastegar@yahoo.com

***Abstract*— In order to correctly share active power across distributed generation (DG) units, this paper uses frequency and active power control in a microgrid consisting of two distributed generation units. The frequency recovery from the microgrid is controlled by an interval type-2 fuzzy logic control with double input, which returns the frequency oscillations in an appropriate range of standard magnitude. The fuzzy controller extends by providing multiple control levels by modifying a customizable membership function that modifies system uncertainty. The output responses for the presence and also absence of the fuzzy control are evaluated to ensure the efficiency of the suggested controller. For examining the influence of output loads on an experimental microgrid, total harmonic distortion has been recognized as a harmonic criterion. Finally, the performance of the suggested architecture is confirmed by the results of the microgrid simulation in island mode.**

***Index Terms* —Microgrid, Interval type 2 fuzzy control, Distributed generation, Frequency control**

I. INTRODUCTION

Despite the many economic and environmental benefits, the presence of distributed generation (DG) and microgrid in the power grid have created additional issues to the power system. Voltage and frequency variations can occur when certain events occur, such as extreme load shifts or power system breakdowns. The intensity and amplitude of these oscillations for unstable microgrids become substantially larger in the event of island operation owing to the lack of backup energy. [1-3].

Furthermore, the dynamics of an island microgrid are significantly quicker than traditional power systems because of the small inertia for the distributed generation in a microgrid and the rapid switching speed of power electronics. As a result, in the event of a system failure or disconnection from the utility grid, an efficient control structure with a quick response is required. [4].

The employing of DGs specifically in microgrids are becoming common procedures in order to prepare reliable active and reactive power for emergency electrical loads. The microgrids have several operating modes such as connecting to main grid and also working in islanding scenario. In non-islanding scenario, the microgrid voltage and frequency parameters are fully controlled by local controller and microgrid distributed generations inject predetermined powers. The microgrid may become an island due to error or pre-planning. In this mode, the function of regulating voltage, frequency and supply of microgrid loads is the responsibility of its DGs. The purpose of this dissertation is to model, control and manage microgrid power. Microgrid dg control can be done in two ways: centralized and decentralized. In the centralized method, there is a need for a central control unit and a communication link between the DGs. In the decentralized method, however, there is no need for communication to control the power of the DGs and the coordination between them. One of the most common methods of decentralized control is the drop method. In this method, only the measured quantities of dg output are used to control them and there is no need for communication between dg units to coordinate them [5-6].

The primary voltage and frequency control loops, as well as the secondary frequency control loop, are subjected to a variety of intelligent approaches. In the principal control approach, known as droop control, a criterion is proposed that employs line characteristics to coordinate voltage and frequency control concurrently under severe load variations [7]. After that, sophisticated techniques like the particle swarm optimization method and fuzzy-neural controllers may be used to optimize the suggested structure. After that, by adding a transient control loop to the proposed technique, the magnitude of DG inverter's current is reduced when the power grid fault occurs, resulting in microgrid transient stability [8]. Because of the application of clean distributed sources, and also low inertia existing in microgrids, even little disturbances might alter basic

characteristics like frequency. In practice, proportional-integral secondary controllers are used to control frequency. Despite the fact that these controllers are inexpensive and simple to use, they are not always the ideal solution for the reasons stated. The use of these controllers and overcoming their challenges may be solved by correcting the control coefficients of these controllers in response to changes in the system. [9-10]. The growing attention in merging clean energies such as PVs and wind turbines in microgrids represents important impulses in terms of control and reliable performance. This article discusses the challenges and important issues in microgrid control. This article categorizes microgrid management procedures in several levels such as: primary, secondary, and also tertiary. In which the primary and secondary levels are related to microgrid performance and third level is related to the cooperated performance for main grid and microgrid.

In order to successfully track self-frequency recovery and active power-sharing, this work offers a unique microgrid control based on interval fuzzy type 2 logic control, which is employed for two DGs. By employing this technology, DGs units may share load fluctuations at a predefined pace and independently restore their output frequency to its nominal value, as well as reduce frequency oscillations.

II. CIRCUIT ARCHITECTURE AND OPERATION PRINCIPLE

The overall system has been illustrated in Fig.1 that is on the basis of [11-12] with some alterations. This system includes a static switch and local controllers (LC), and also linear loads, two DGs system, a central controller. At first the main grid is vividly seen that is connected to the microgrid and the linear loads feed from both sources. Then, the disconnection of microgrid after a while is occurred from the main grid and static switch is opened and as a result of that, microgrid must works independently in islanded mode. It can be noticed that the fundamental task of central controller is adjusting the voltage and frequency for microgrid at standard magnitude and also work efficiently the microgrid in all the conditions. However, local controller has been employed for two DGs to control the output power of each DGs autonomously.

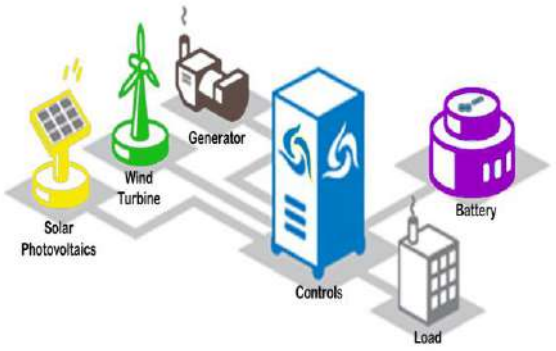


Fig.1. Structure of microgrid consists of two DGs.

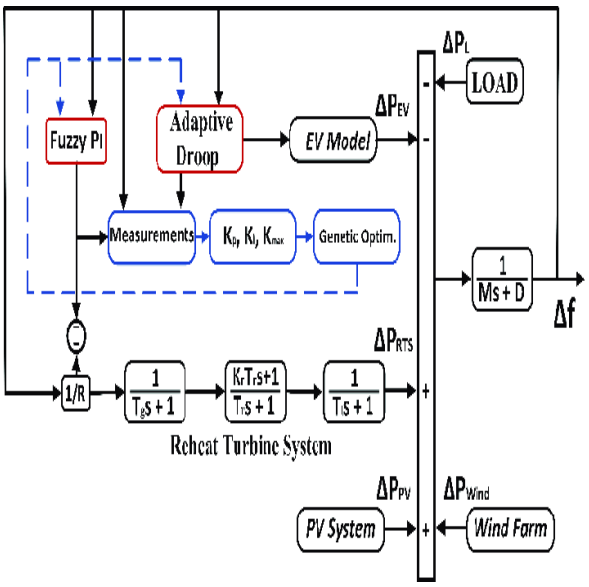


Fig.2. control structure for DG system of microgrid.

The DG system control structure in microgrid is shown in Fig .2 which is composed of two sections. The first section is frequency and active power droop control that becomes responsible to adjust main frequency at the primary side of the power grid. In order to support that self-frequency control is added in structure in order to compensate the extra alteration and tracking the reference value of the system. Frequency control and active power sharing among DGs will be concentrating in this paper. Moreover, the voltage control loop is added by adjusting the reactive power to keep the amplitude of voltage in desired range and reference voltage is obtained from equation (1):

$$V_{i,ref} = V_o + n_i(Q_{i,dis} - Q_i) \qquad (1)$$

Vo is defined as reference voltage and also n_i equals to droop coefficient.

A:Droop control

The traditional f-P droop control is used in basic control of active power sharing. The frequency fi of the output may be written as:

$$f_i = f_{nom} + m_i(P_{i,dis} - P_i) \qquad (2)$$

The variation of the active power from the distributed generation (given by the Central Controller), that is related to f-P droop coefficient mi, was used to establish the reference output frequency.

B:Self-Frequency Recovery Control

At first, main goal of self-frequency recovery strategy is allocation of values which are necessary for frequency recovery between distributed generation units participating in active power sharing with f-P droop control in a predefined ratio. Due to self-frequency recovery control, the frequency of the i_{th} DG unit can be restored as follows:

$$\Delta f_{i,res} = k_f \int (f_{nom} - f_i)dt \qquad (3)$$

In which k_f is similar for distributed generation units, implying that frequency restoration responsibility is distributed evenly across the DGs. Furthermore, for better performance of this segment, interval type 2 fuzzy control is used as a replacement of PI control to ameliorate the behavior of the system which will be elaborated in the following.

III. INTERVAL TYPE-2 FUZZY CONTROL

Fuzzy logic control is suitable mathematical structure which assesses input data in terms of logical variables having continuous values between 0 and 1. A fuzzy control system is one that is based on fuzzy logic. Fuzzy logic controllers convert expert knowledge in automatic control techniques for linguistic strategies into control.

A fuzzy controller is used to compensate for the PQ issue. The phases in a fuzzy controller include fuzzification, decision making, and defuzzification. Fuzzification is the process of transforming a sharp value into a fuzzy one. To achieve the fuzzification process, which has no set procedure, many types of fuzzifiers are utilized. Fuzzy sets exist in a range of shapes and sizes, including triangular, trapezoidal, and other geometric shapes. In the decision-making process, the fuzzified output is applied to a set of rules. The bias voltage generator's input is selected from FIS using fuzzy criteria. The fuzzified computed voltage is then calculated using the defuzzification process (Vdc). The fuzzy logic controller is a complex controller that can handle several converters in different scenarios.

Then, interval Type-2 Fuzzy Sets are specified as a mathematical method to fuzziness which considers uncertainty for model created elaborately, on the basis of the fundamental notions of Fuzzy Sets. In recent years, several advancements in this method have been made. An IT2FS is expressed mathematically is written in Eq. 4.

$$\tilde{A} = \{\{(x,u),1) \mid \forall x \in X, \forall u \in J_x \subseteq [0,1]\} \qquad (4)$$

In which the X parameter is the primary domain, which reflects the fuzzy set's membership degree, and also the Jx parameter is the secondary domain, which is always equal to 1. The top member ship function (MF) and lower membership function are two T1 MFs, and Footprint of Uncertainty (FOU) [8] is area which is among them, as shown in Figure 3.

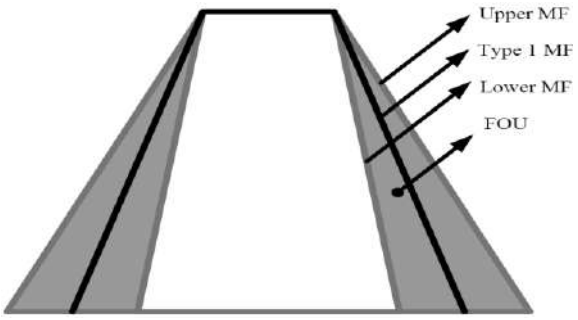


Fig.3. Interval Type 2 Membership Function.

Also, construction of a FIS on the basis of Interval Type 2 FL, on the other hand, is depicted in Figure 4.

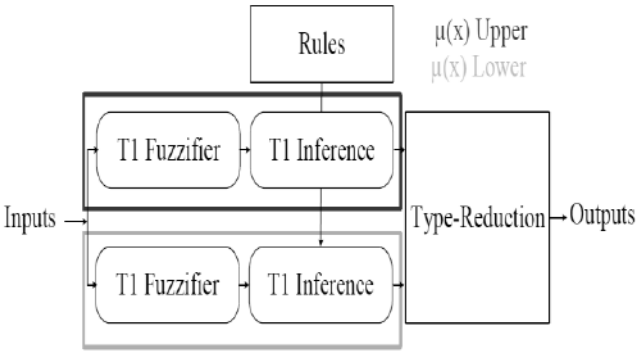


Fig.4 Structure of mamdani IT2 FIS.

Then, the trapezoidal membership function as written in equation 5 is considered for control strategy.

$$IT\ 2MF = \begin{cases} trapmf\ (x, [a - \frac{u}{2}, b, c, d + \frac{u}{2}]) \\ trapmf\ (x, [a + \frac{u}{2}, b, c, d - \frac{u}{2}]) \end{cases} \qquad (5)$$

The design fuzzy control addresses the lower and upper MF, and u is the suggested parameter to adjust FOU variable of the IT2MF. Figure 3 depicts graphical depiction of Equation 5.

By the way, improving the performance of converter fuzzy logic control has been used for DC link voltage. The triangle membership function has been employed for the fuzzy logic and also Table 1 shows the rule based for the converter. It can be found that e and de are error and derivative of error signal respectively for controlling the output voltage. PB, PM and PS and PVS are equal to positive big, positive medium, positive small and positive very small respectively and ZR equals to zero.

Table1-Rule based of fuzzy control system.

NB	NM	NS	NVS	ZR	PVS	PS	PM	PB	e / de
ZR	PVS	PS	PM	PB	PB	PB	PB	PB	PB
NVS	ZR	PVS	PS	PM	PB	PB	PB	PB	PM
NS	NVS	ZR	PVS	PS	PM	PB	PB	PB	PS
NM	NS	NVS	ZR	PVS	PS	PM	PB	PB	PVS
NB	NM	NS	NVS	ZR	PVS	PS	PM	PB	ZR
NB	NB	NM	NS	NVS	ZR	PVS	PS	PM	NVS
NB	NB	NB	NM	NS	NVS	ZR	PVS	PS	NS
NB	NB	NB	NB	NM	NS	NVS	ZR	PVS	NM
NB	NB	NB	NB	NB	NM	NS	NVS	ZR	NB

IV. Simulation Results

The issue of frequency control in microgrids is important despite the scattered production resources along with energy storage equipment. Due to changes in the value of the distributed generation (such as variable wind turbine system and also photovoltaic system) and due to changes in load consumption power, the microgrid frequency has changed, which requires the design of controllers to improve the frequency. In the design of controllers, the dynamic response of each manufacturer is modeled and simulated and used in the microgrid model. To eliminate frequency fluctuations, it is necessary to make a balance among generation and consumption for output loads. To do this, the Drop control strategy has been used based on the nominal value of the producers and the capability of each of them. To improve the performance of Drop controller, classical controller and fuzzy controller have been used, each of which has its own characteristics and has improved control in the system and reduced frequency fluctuations in the network. Due to the uncertainty in the microgrid structure and the lack of precise adjustment of controller parameters, optimization has been used to achieve the goal of minimizing frequency fluctuations.

In this section, various scenarios for the concerned microgrid, as illustrated in Fig. 5, are studied deeply and investigated for validating efficiency of recommended self-frequency recovery control. Table 2 also listed the microgrid's characteristics. The influence of electrical load change and islanding on the microgrid's dynamic behavior was investigated in transient scenarios (such as opening the static switch at t = 0. 5 s).

Input parameters of microgrid is the AC voltage source and also variables of the renewable energies such as temperature and radiation for PV system and wind speed for wind turbine and output parameters of the microgrid is the voltage and frequency and reactive power and active power which are achieved and investigated in output results deeply.

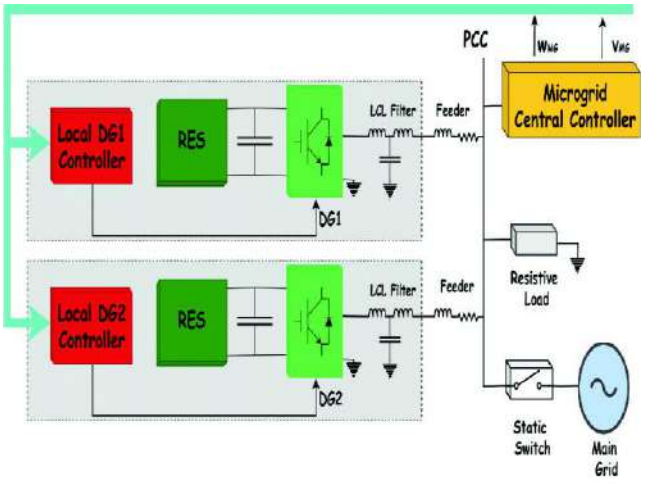


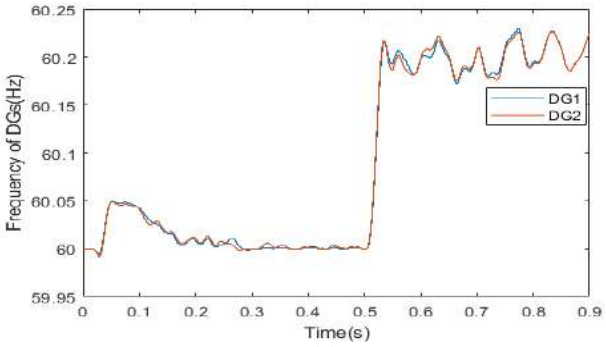
Fig. 5. Structure of the microgrid system

Table 2. Properties of microgrid.

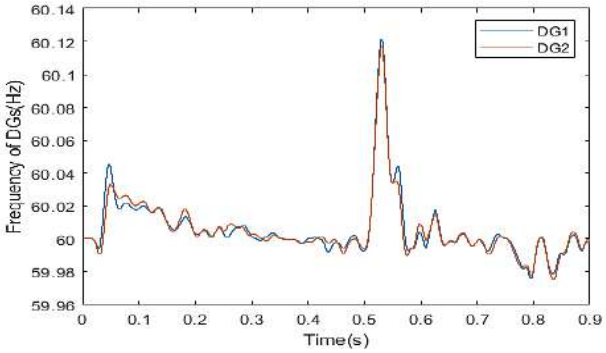
Parameter	Value
Rated frequency	60Hz
Rated voltage	13.8kV
LC inductance	L=5mH
LC capacitance	C=50μF
Static switch	Initially closed, and then opened at 0.5s.
Active power dispatched value	1.3MW
Reactive power dispatched value	0MVAR
Q-V droop coefficient	$K_1=K_2=K_3=0.01$
P-f droop coefficient	$M_1=M_2=M_3=0.02$
Linear loads	$L_1=L_2=0.5$MW

It can be seen in Fig. 6 (b) that using of interval type 2 fuzzy logic control, the microgrid is prone to restore from the frequency oscillations and also reaching to its standard value (i.e., f_0 = 60 Hz) with respect to the microgrid code requirements in a better way compared to PI control in Fig . 6 (a) . In addition, it could be witnessed that the response of frequency for type 2 fuzzy logic control experiences small overshoot while the static switch is opened and the microgrid operates in islanding mode. It is observed that by adopting interval type 2 fuzzy control, effectively output frequency can

be investigated for various DGs are swiftly arranged in a short range of standard frequency, and the frequency variation shown in the P–f control droop is effectively adjusted.



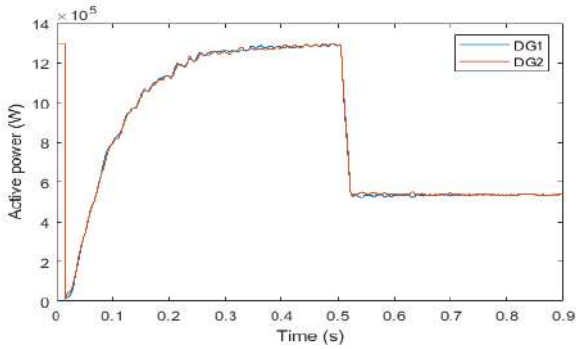
(a)



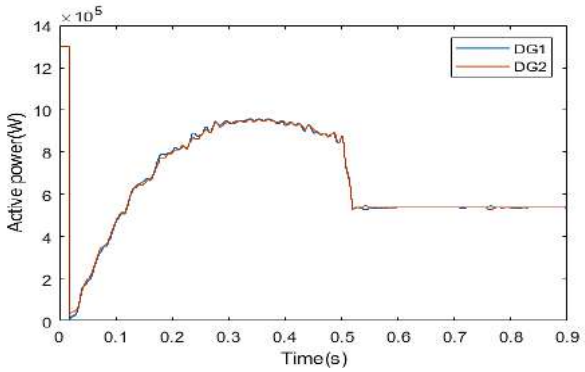
(b)

Fig. 6. Output frequency of DGs in microgrid (a)PI control (b) interval type 2 fuzzy control.

Figures 7 and 8 show the effectiveness of managing the active and reactive power for DGs employing two alternative methods: PI control and interval type 2 control. Due to the associated parameters for P–f and also Q–V droop are regarded identical for every unit, the required active power injected from the power grid is evenly distributed between the Distributed Generation units after opening static switch at 0.5 s.



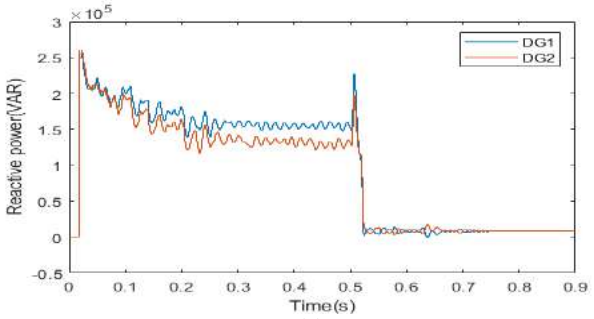
(a)



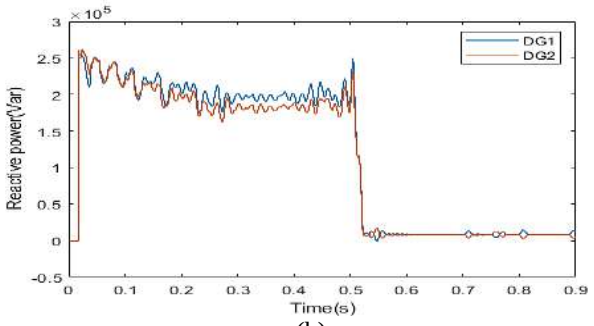
(b)

Fig. 7. Active power of DGs in microgrid (a)PI control (b) interval type 2 fuzzy control.

In Figure 9 suitable performance for voltage bus of DGs in microgrid is presented. After being in islanding mode for microgrid the amplitude of buses is differed and experiences a small increasing in the value. The oscillations of buses voltage reduced significantly by proposed controller compared to conventional PI controller.



(a)

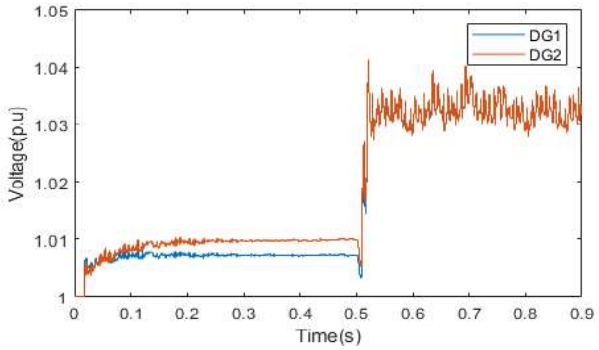


(b)

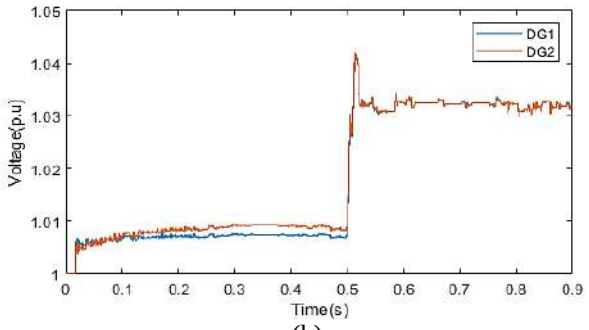
Fig. 8. Reactive power of DGs in microgrid (a) PI control (b) interval type 2 fuzzy control.

Despite the fact that it may be preferable to distributed parameter for frequency restored between Distributed

Generation units by using only certain specific DG units (such as controlling of frequency units) in view of residual power of the frequency control units, generation, the optimum ratio of active power exchanged between Distributed Generation units must be calculated on the basis of a proper purpose. For example, generation cost of DG units, and also losing sensitivity might be represented by this ratio. Furthermore, the suggested control system must be adjusted in order to be implemented in low-voltage networks. Proposed technique modifies DG frequency nominal value; although, because the network reactance is larger than the resistance, this control system can only be used in high-voltage or medium voltage networks with well-decoupled frequency and reactive power. Moreover, relationship among reactive and active power is accurate to use.



(a)



(b)

Fig. 9. Bus voltage of DGs in microgrid (a) PI control (b) interval type 2 fuzzy control.

V. CONCLUSION

Nowadays, increasing the reliability of customer service and reducing congestion and losses in transmission and distribution lines, distributed and renewable energy sources, are new and suitable options that have been introduced in the modern power system. Actually, the benefits of renewable energy sources, leading to planning to use them more and more recently. In the electricity industry and the growing expansion of microgrids in most countries of the world. Microgrids are small power grids composed of several renewable energy sources and local loads. Microgrids are normally connected to one of the bursars of the distribution network; But in an emergency, in the event of a major disturbance, they are disconnected from the power grid and fed some loads (important loads). Due to significant changes in wind speed, solar intensity, changes and load fluctuations are one of the important issues in microgrids separate from the main network, control and damping of frequency and power fluctuations. This is also the case in the time periods corresponding to the operation and management, which aims to dampen power fluctuations. A unique interval type-2 fuzzy logic control mechanism for DG units has been devised in order to combine precise self-frequency recovery and also active power sharing in an islanded microgrid. Due to their small inertia, islanded microgrid is prone to frequency disruptions, necessitating frequency modification. Traditionally, frequency restoration will be handled by secondary frequency control units, with active power sharing Distributed Generation units by local control individually. To adapt for load shifts, special components (such as frequency control units) are required, which may allow them to hit their performance limit sooner and hence exponentially raise the cost of generating. Furthermore, whether frequency oscillation is high, the frequency control units' potential might be lost, specifically for a tiny isolated power grid like an islanded microgrid.

References

[1] H. Xin *et al*., "Control of island AC microgrids using a fully distributed approach," *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp. 943–945, Mar. 2015.

[2] J. He, Y. W. Li, J. M. Guerrero, F. Blaabjerg, and J. C. Vasquez, "An islanding microgrid power sharing approach using enhanced virtual impedance control scheme," *IEEE Trans. Power Electron*, vol. 28, no. 11, pp. 5272–5282, Nov. 2013.

[3] Zia, M.F., Benbouzid, M., Elbouchikhi, E., Muyeen, S.M., Techato, K. and Guerrero, J.M., 2020. Microgrid transactive energy: Review, architectures, distributed ledger technologies, and market analysis. *Ieee Access*, *8*, pp.19410-19432.

[4] Moghadam, H.M., Khooban, M.H., Dragicevic, T. and Masoudian, A., 2018, November. Using Interval Type2 Fuzzy Controller in Ship Power Systems in Presence of Pulsed Power Loads. In *2018 IEEE International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles & International Transportation Electrification Conference (ESARS-ITEC)* (pp. 1-6). IEEE.

[5] Mohammadi Moghadam, H., Gheisarnejad, M., Yalsavar, M., Foroozan, H. and Khooban, M.H., 2021. A Novel Nonsingular Terminal Sliding Mode Control-Based Double Interval Type-2 Fuzzy Systems: Real-Time Implementation. *Inventions*, *6*(2), p.40.

[6] Zarepour, A., Rajaei, A., Mohammadi-Moghadam, H. and Shahparasti, M., 2021. A High Gain AC-DC Rectifier Based on Current-Fed Cockcroft Walton Voltage Multiplier for Motor Drive Applications. *Sustainability*, *13*(21), p.12317.

[7] Gheisarnejad, M., Mohammadi-Moghadam, H., Boudjadar, J. and Khooban, M.H., 2019. Active power sharing and frequency recovery control in an islanded microgrid with nonlinear load and nondispatchable DG. *IEEE Systems Journal*, *14*(1), pp.1058-1068.

[8] Mohammadzadeh, A. and Kayacan, E., 2020. A novel fractional-order type-2 fuzzy control method for online frequency regulation in ac microgrid. *Engineering Applications of Artificial Intelligence*, *90*, p.103483.

[9] Mohammadi Moghadam, H., Mohammadzadeh, A., Hadjiaghaie Vafaie, R., Tavoosi, J. and Khooban, M.H., 2021. A type-2 fuzzy control for active/reactive power control and energy storage management. *Transactions of the Institute of Measurement and Control*, p.01423312211048038.

[10] Sabahi, K., Tavan, M. and Hajizadeh, A., 2021. Adaptive type-2 fuzzy PID controller for LFC in AC microgrid. *Soft Computing*, *25*(11), pp.7423-7434.

[11] Moghadam, H.M., Gheisarnejad, M., Esfahani, Z. and Khooban, M.H., 2020. A Novel Supervised Control Strategy for Interconnected DFIG-Based Wind Turbine Systems: MiL Validations. *IEEE Transactions on Emerging Topics in Computational Intelligence*

[12] Cao, B., Dong, W., Lv, Z., Gu, Y., Singh, S. and Kumar, P., 2020. Hybrid microgrid many-objective sizing optimization with fuzzy decision. *IEEE Transactions on Fuzzy Systems*, *28*(11), pp.2702-2710.

[13] Y.-S. Kim, E.-S. Kim, and S.-I. Moon, "Distributed generation control method for active power sharing and self-frequency recovery in an islanded microgrid," *IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 544–551, Jan. 2017.

[14] Sahoo, G., Sahu, R.K., Samal, N.R. and Panda, S., 2021. Analysis of type-2 fuzzy fractional-order PD-PI controller for frequency stabilisation of the micro-grid system with real-time simulation. *International Journal of Sustainable Energy*, pp.1-22.

[15] Afrakhte, H. and Bayat, P., 2020. A self-evolving type-2 fuzzy energy management strategy for multi-microgrid systems. *Computers & Electrical Engineering*, *85*, p.106702.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

A Novel Ensemble Feature Selection Method through Type I Fuzzy

Nazanin Zahra Joodaki[1], Mohammad Bagher Dowlatshahi[2], Mehdi Joodaki[3]

[1] *Department of Computer Engineering, Lorestan University, Khorramabad, Iran*, joodaki.nz@fe.lu.ac.ir
[2] *Department of Computer Engineering, Lorestan University, Khorramabad, Iran*, dowlatshahi.mb@lu.ac.ir
[3] *Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan, Iran*, mehdi.joodaki@ec.iut.ac.ir

***Abstract*— These days, one of the needed methods in machine learning is feature selection. In other words, in this manner, the most fitting features are picked. Nevertheless, there are various feature selection methods, getting the most suitable features still is a complex problem. Lately, applying several feature selection methods rather than a unique feature selection method is more efficient. In this paper, a new ensemble feature selection method based upon fuzzy Type-I named EFSF is presented. First, three different individual feature selection methods are applied to determine the rank of features separately. Next, Type-I fuzzy handles feature selections' uncertainty and decrease noise to give each feature the best rank. To validate the act of EFSF, it is compared with some ensemble methods and several advanced feature selection methods. EFSF is assessed based on Accuracy, Precision and Recall, metrics. The outcomes verify that the EFSF is better than its competitors. The source code of EFSF is here.**

***Index Terms* —Feature selection; Type-I fuzzy; Ensemble feature selection; High-dimensional datasets**

I. INTRODUCTION

Determining the most essential and suitable features is a necessary manner in all high dimensional datasets. Many systems, because of the inadequacy of enough hardware, cannot analyze high-dimensional data. As a result, one of the main approaches to tackle high-dimensional data is dimensionality reduction. The purpose of dimensionality reduction is applied in two classes of feature selection and feature extraction. In feature extraction, the features are mixed and given into new space, while feature selection, trying to overcome the irrelevant features and select a fit group of features [1-6]. Some features take a heavy toll on the learning process. Therefore, maintaining an appropriate group of features enhances the model's efficiency and reduces memory and computation costs [7-9]. The three main groups of feature selection methods are wrapper, filter, and embedded [7]. The filter method takes features while neglectful of the learning algorithm [10]. The usual measurements applied in this process are based on statistical metrics such as distance [2]. Here, several filter methods are presented.

For instance, the Correlation [11] decide extremely correlated features to the target variable. The Fisher score method [12], as a supervised method, uses the maximum likelihood. The Chi-square [13] select the proper features based on chi-square metrics and distribution. The ReliefF [14] determines restrictive dependencies between features. The RFS [15] gains projection between the target and features in terms of the R-value. The MIC [16] analyses the relationships of features in terms of the coefficient between them.

Some feature selection techniques imply a wrong rank for their inputs that are diverse from the genuine ranks and bring about noise. Therefore, a mixture method should be able to manage noise. This paper aims at introducing a novel ensemble approach. This paper's primary offering is a novel combination method based on Type-I fuzzy logic that manages diverse feature selections' uncertainty and overwhelms noise to obtain the best rank for each feature. In the same vein, this paper uses the majority rule [4], which indicates that a rank far from the other ranks is probably far from the actual rank.

Here, three feature selections are employed separately for ranking features and then developing a voter based on Type-I fuzzy logic to judge between ideas of various feature selection methods. The proposed method is compared with some similar procedures and base feature selection methods on several real-world datasets. The outcomes prove that the proposed method produces superb results compared to the other methods in terms of all metrics.

This paper's construction is arranged as follows: In Section II, some related works are reviewed. EFSF is presented in Section III, and the productions are discussed in Section IV. Lastly, in Section V, a consequence is given.

II. Related Work

Most of the current datasets are huge and have an immense number of features, so it is essential to reduce select a part of features to get the best result [17-20]. Using ensemble feature selection methods have become widespread. For instance, EFSPF [1] firstly applies three feature selection methods and then by fuzzy logic, a reliable weighted graph is built. Finally, the Weighted Page Rank algorithm is employed to rank the features. In [2], to begin with, the correlation between the target feature and other features in multi-label datasets are calculated. Then, a weighted graph is created and the proper features are found by running Page Rank. MLACO [7], by using Ant Colony Optimization (ACO) builds an ensemble feature selection method. In other words, in this method, ACO is applied in multi-label datasets to select the most relevant features with the lowest redundancy with themselves and most maximum relevancy with the target variable. In PEFS [6], the bi-objective optimization problem and the crowding distance are engaged to determine the best group of features. In [10], a collection of simple and advanced strategies comprising Max, Mean, and Min are used.

III. Proposed Method

In most cases, each dataset consists of a vector of features; they influence the classification act differently. Some features may boost the classification performance and vice versa. There are many feature selection methods with their designs to rank or take an optimal subset of features. For gaining more reliable results, employing multiple feature selection methods can be more effective. On the other hand, if we consider multiple feature selection methods, it is efficient, helpful, and have a dominant influence on the final result. However, feature selection methods may propose an inappropriate rank for some features and bring noise. So, we need to take a robust approach for blending efficiently and reaching a settlement between feature selection methods' opinions.

The problem of ensemble feature selection is a merging or combination problem, and there are many strategies for mixing feature selection opinions, and some methods like min, max, mean, and medium are simple. The designs based on fuzzy sense are proper for melding distinguishable methods and eliminating noise. In other words, fuzzy logic gives more flexibility for the uncertainty of methods when each feature's ranks for various methods are uncertain or noise. Consequently, in this paper, a new ensemble feature selection method that supports the majority rule (means a rank that is far from the other ranks is probably far from the actual rank) has been created based on fuzzy logic to discover the best rank for each feature and handle noise.

Here, to begin with, three feature selection methods, inclusive t-test, RFS and Maximal Information Coefficient, are applied. Next, the contrast within normalized inputs in the scale (0,1) is accounted. Next, the fuzzy voter for refereeing is set as follows.

As stated above, after running three feature selection methods, we have three various ranks for each feature. As a result, we aim at developing a voter to judge in these ranks.

Imagine we have three ranks R1, R2, and R3 as calculated ranks for feature *X*. The Type-I fuzzy voter in the first stage estimate the distance between pair two calculated ranks (*D12*=|*R1-R2*|, D13=|*R1-R3*|, *D23*=|*R2-R3*|). After that, Small, Medium, and Large as membership functions (Fig. 1) are used to fuzzify the distances. Next, for estimating the agreement between them, Vlow, Low, Medium, High, and Vhigh membership functions are employed (Fig. 2).

The parameters and rules (Table I-Table III) for the mentioned membership function in EFSPF are used here that they have been designed by an expert. For example, consider this rule:

"If D12 is Small and D13 is Large, and D23 is Large, then output is Low"

This rule shows one of the used rules in this paper. In other words, with these rules, the voter assesses the contrast between various methods. It shows if the contrasts between methods 1 and 2, 1 and 3, 2 and 3 are Small, Large and Large, respectively, then the rule's output is Low.

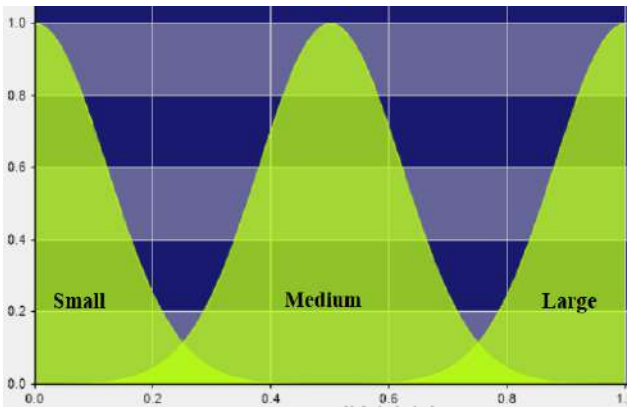


Figure 1. Membership functions of fuzzifying [1].

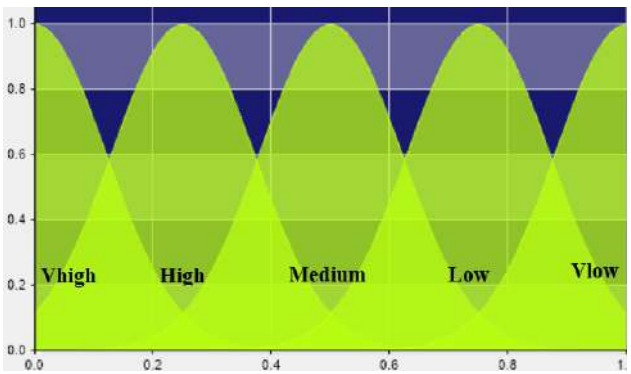


Figure 2. Membership functions of agreement [1].

Table I. Several linguistic rules [1].

D12	*D13*	*D23*	Output
Small	Small	Small	Vhigh
Small	Medium	Small	Vhigh
Small	Large	Small	High
Large	Large	Large	Vlow
Small	Large	Large	Low
Small	Medium	Medium	Medium
Medium	Large	Large	Vlow
Medium	Medium	Medium	Medium
Medium	Large	Medium	Medium

Table II. Fuzzification's parameters [1].

Small		Medium		Large	
σ	M	σ	m	σ	m
0.12	0	0.12	0.5	0.12	1

Table III. Agreement's parameters [1].

Vhigh		High		Medium		Low		Vlow	
σ	m	σ	m	σ	m	σ	m	σ	m
0.12	0	012	0.25	0.12	0.5	0.12	0.75	0.12	1

Finally, the centroid (C_j) of N (quantity of fuzzy sets) fuzzy sets (B_j) is calculated. Also, the membership value (μ_{B_j}) of each fuzzy set is computed. At the end, the final vote for the X the Eq. (1) is used:

$$W_X = \frac{\sum_{j=1}^{N} \mu_{B_j} C_j}{\sum_{j=1}^{N} \mu_{B_j}} \quad (1)$$

The primary voter system is formed by utilizing the Fuzzification membership functions to convert distances into fuzzy lexical variables. The agreement membership functions are implemented to estimate the settlement within methods.

In this voter, the Gaussian functions because of their high accuracy for various problems are selected. Because judged ranks by various methods are normalized in the range (0,1), the parameters for fixing up the Gaussian process are developed through Tables II and III.

IV. EXPERIMENTS

Six datasets are used to analyze the proposed method with the most advanced ensemble methods and based methods (Table IV). To explain the performance of the proposed method, it has been compared with various advanced methods. In other words, plurality vote, Borda voting, Weighted Borda and a group of based methods have been compared with the proposed method in terms of accuracy (Figs 3-14). Accuracy is the rate of correct classification.

In other words, for assessing the act of classification, the KNN algorithm is applied. Any dataset is split erratically into 80 per cent for the train dataset and 20 per cent for the test dataset. Next, EFSF and its rivals use the 100 high selected ranks of inputs (each dataset). Then, for per dataset, KNN is implemented 21 rounds. In the first step, one feature is picked. Then, five features are added in each round in order to all 100 high ranks of features are picked.

Our research applied several datasets to demonstrate our method's production in Accuracy, Precision, Recall, and F1-score measurements. Accuracy is the rate of accurately matched and F1-score assessed with precision and recall.

The precision is the amount of factual positive cases intersected by the total of all cases as positive, and recall is the total of factual positive cases intersected by the whole quantity of factual positive cases. Accordingly, the mentioned metrics are computed by below formulas:

$$Presision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1-score = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}} \quad (4)$$

Table IV. Information about the used datasets.

Dataset	Instance	Feature	Class
Jaffe **[21]**	213	676	10
ORL **[22]**	400	1024	40
Yale **[23]**	165	1024	15
WarpPIE10P **[24]**	210	2421	10
Lung **[25]**	203	3312	5
GLIOMA **[10]**	50	4434	4

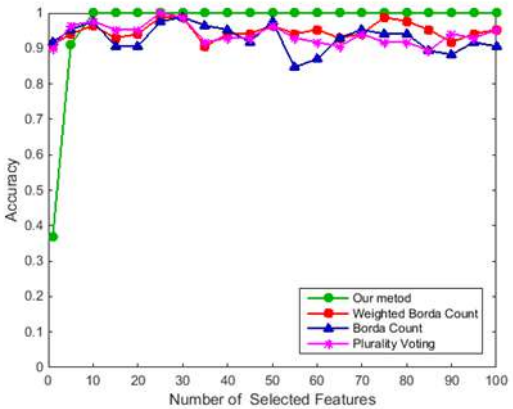


Figure 3. Outcomes of KNN based on Accuracy on Jaffe (Ensemble methods).

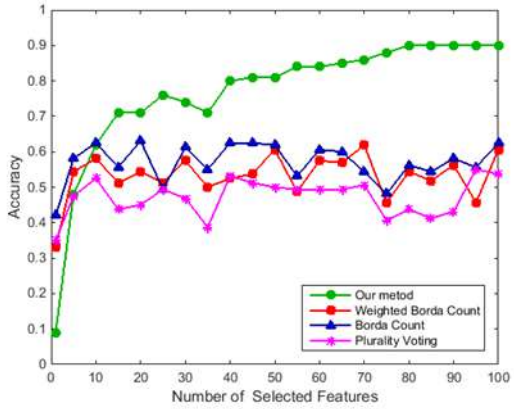


Figure 4. Outcomes of KNN based on Accuracy on ORL (Ensemble methods).

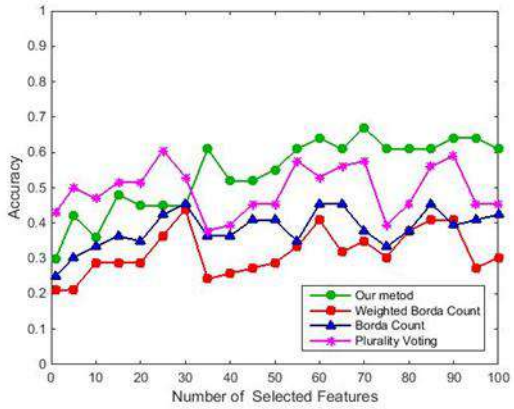


Figure 5. Outcomes of KNN based on Accuracy on Yale (Ensemble methods).

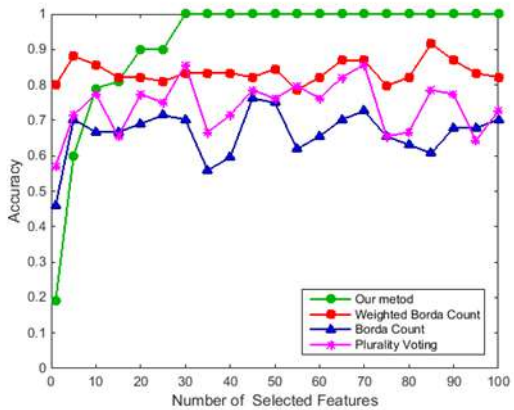


Figure 6. Outcomes of KNN based on Accuracy on WarpPIE10P (Ensemble methods).

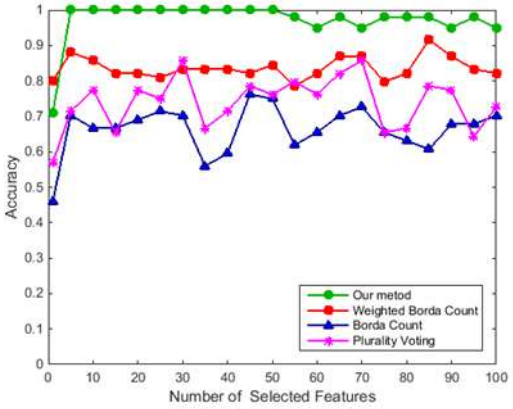


Figure 7. Outcomes of KNN based on Accuracy on Lung (Ensemble methods).

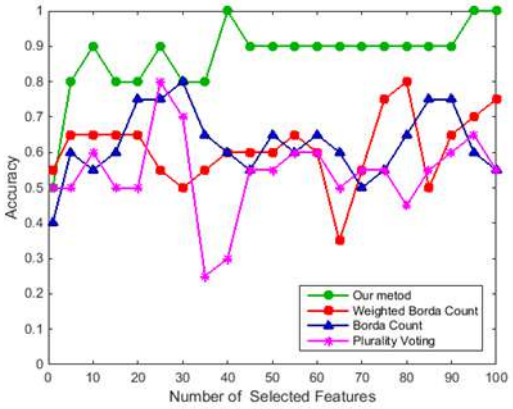


Figure 8. Outcomes of KNN based on Accuracy on GLIOMA (Ensemble methods).

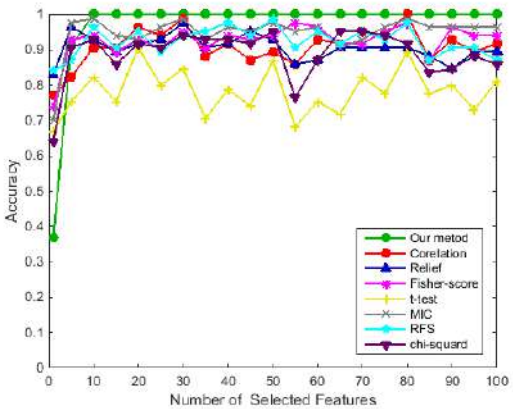


Figure 9. Outcomes of KNN based on Accuracy on Jaffe (Based methods).

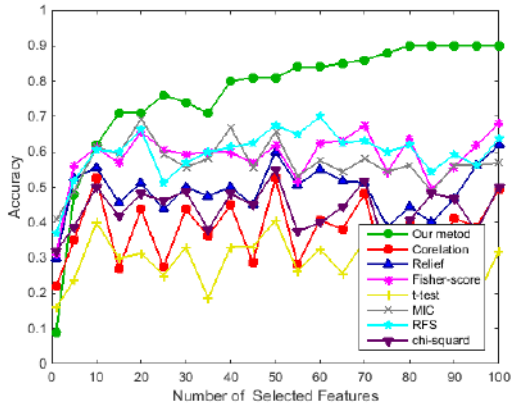


Figure 10. Outcomes of KNN based on Accuracy on ORL (Based methods).

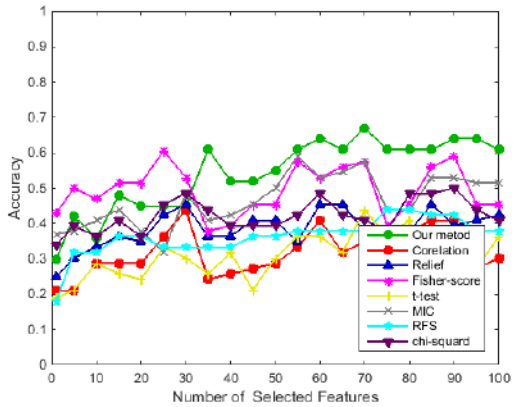


Figure 11. Outcomes of KNN based on Accuracy on Yale (Based methods).

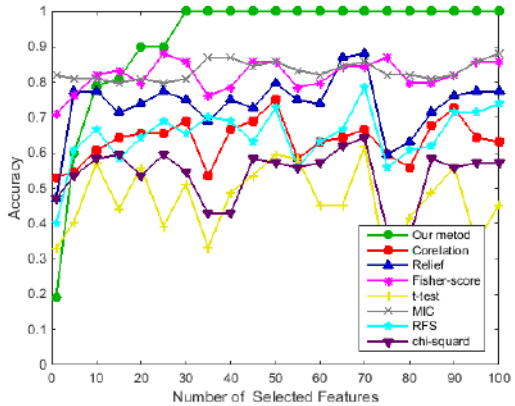


Figure 12. Outcomes of KNN based on Accuracy on WarpPIE10P (Based methods).

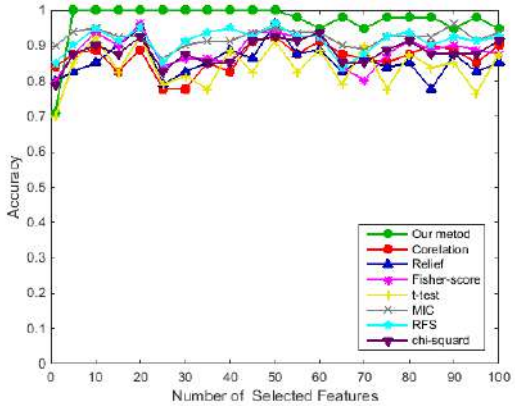


Figure 13. Outcomes of KNN based on Accuracy on Lung (Based methods).

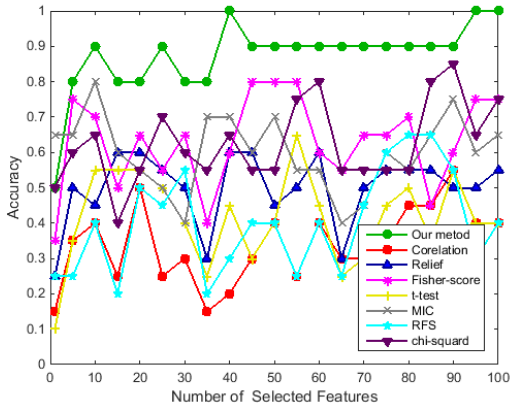


Figure 14. Outcomes of KNN based on Accuracy on GLIOMA (Based methods).

It can be seen from Figs. 3-14 and other Tables (Precision, Recall) in appendix A that EFSF has earned higher Accuracy, Precision, Recall, F1-score and better result than other methods. This success is due to adopting a suitable ensemble feature selection method. In other words, the compared ensemble feature selection methods do not consider wrong ranks, and they combine all ranks with the same weight to reach a final rank. On the other hand, based methods work just on one rank, and it is more likely that the rank is wrong. In marked contrast, the proposed method combined several ranks and tried to lessen the effect of wrong ranks with its strategy. The voter through fuzzy is a soft and effective method that regards uncertainty between various strategies with different ideas. That is to say, the ideas of various methods are not similar; as a result, methods regard a wrong rank and yield noise in ranks.

Consequently, when we want to fusion all opinions, the noise takes a toll on the model performance. Therefore, the proposed method has tried to defeat noise influence; it is evident that it has had the best act over other methods because it has superbly dealt with noise.

V. Conclusion

Running multiple feature selection methods produces better results than employing only a particular feature selection method. However, fusion multiple feature selection methods can plummet the model performance because of noise in ranks. Here, we have presented a new manner for ensemble feature selection that three approaches are run separately to rank features. After that, the contrast between forecasted ranks is computed.

Then, a fuzzy voter is set, and the built voter to judge within ideas of FS methods and lessen the influence noise is implemented. The outcomes show the built voter referee notably within ideas of feature selection methods. The result will be much better in using fuzzy Type-II to set the membership function for future work.

References

[1] M. Joodaki, M. B. Dowlatshahi, and N. Z. Joodaki, "An ensemble feature selection algorithm based on PageRank centrality and fuzzy logic," *Knowledge-Based Systems*, p. 107538, 2021.

[2] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-pour, "MGFS: A multi-label graph-based feature selection algorithm via PageRank centrality," *Expert Systems with Applications*, vol. 142, p. 113024, 2020.

[3] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-pour, "MFS-MCDM: Multi-label feature selection using multi-criteria decision making," *Knowledge-Based Systems*, vol. 206, p. 106365, 2020.

[4] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-Pour, "A bipartite matching-based feature selection for multi-label learning," *International Journal of Machine Learning and Cybernetics*, vol. 12, no. 2, pp. 459-475, 2021.

[5] A. Hashemi and M. B. Dowlatshahi, "MLCR: a fast multi-label feature selection method based on K-means and L2-norm," in *2020 25th International Computer Conference, Computer Society of Iran (CSICC)*, 2020, pp. 1-7: IEEE.

[6] A. Hashemi, M. Bagher Dowlatshahi, and H. Nezamabadi-pour, "A Pareto-based ensemble of feature selection algorithms," *Expert Systems with Applications*, p. 115130, 2021/04/30/ 2021.

[7] M. Paniri, M. B. Dowlatshahi, and H. Nezamabadi-pour, "MLACO: A multi-label feature selection algorithm based on ant colony optimization," *Knowledge-Based Systems*, vol. 192, p. 105285, 2020.

[8] H. Bayati, M. B. Dowlatshahi, and M. Paniri, "MLPSO: a filter multi-label feature selection based on particle swarm optimization," in *2020 25th International Computer Conference, Computer Society of Iran (CSICC)*, 2020, pp. 1-6: IEEE.

[9] H. Bayati, M. B. Dowlatshahi, and M. Paniri, "Multi-label feature selection based on competitive swarm optimization," *Journal of Soft Computing and Information Technology*, vol. 9, no. 4, pp. 56-69, 2020.

[10] P. Drotár, M. Gazda, and L. Vokorokos, "Ensemble feature selection using election methods and ranker clustering," *Information Sciences*, vol. 480, pp. 365-380, 2019.

[11] L. Yu and H. Liu, "Feature selection for high-dimensional data: A fast correlation-based filter solution," in *Proceedings of the 20th international conference on machine learning (ICML-03)*, 2003, pp. 856-863.

[12] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*. John Wiley & Sons, 2012.

[13] I. H. Witten and E. Frank, "Data mining: practical machine learning tools and techniques with Java implementations," *Acm Sigmod Record*, vol. 31, no. 1, pp. 76-77, 2002.

[14] M. Robnik-Šikonja and I. Kononenko, "Theoretical and empirical analysis of ReliefF and RReliefF," *Machine learning*, vol. 53, no. 1-2, pp. 23-69, 2003.

[15] J. Lee, N. Batnyam, and S. Oh, "RFS: Efficient feature selection method based on R-value," *Computers in biology and medicine*, vol. 43, no. 2, pp. 91-99, 2013.

[16] E. F. Combarro, E. Montanes, I. Diaz, J. Ranilla, and R. Mones, "Introducing a family of linear measures for feature selection in text categorization," *IEEE transactions on Knowledge and Data Engineering*, vol. 17, no. 9, pp. 1223-1232, 2005.

[17] S. Samiei, M. Joodaki, and N. Ghadiri, "A Scalable Pattern Mining Method Using Apache Spark Platform," in *2021 7th International Conference on Web Research (ICWR)*, 2021, pp. 114-118: IEEE.

[18] M. Joodaki, N. Ghadiri, Z. Maleki, and M. L. Shahreza, "A scalable random walk with restart on heterogeneous networks with Apache Spark for ranking disease-related genes through type-II fuzzy data fusion," *Journal of Biomedical Informatics*, p. 103688, 2021.

[19] M. Joodaki, N. Ghadiri, and A. H. Atashkar, "Protein complex detection from PPI networks on Apache Spark," in *2017 9th International Conference on Information and Knowledge Technology (IKT)*, 2017, pp. 111-115: IEEE.

[20] S. V. Moravvej, M. Joodaki, M. J. M. Kahaki, and M. S. Sartakhti, "A method Based on an Attention Mechanism to Measure the Similarity of two Sentences," in *2021 7th International Conference on Web Research (ICWR)*, 2021, pp. 238-242: IEEE.

[21] M. J. Lyons, J. Budynek, and S. Akamatsu, "Automatic classification of single facial images," *IEEE transactions on pattern analysis and machine intelligence*, vol. 21, no. 12, pp. 1357-1362, 1999.

[22] F. S. Samaria and A. C. Harter, "Parameterisation of a stochastic model for human face identification," in *Proceedings of 1994 IEEE workshop on applications of computer vision*, 1994, pp. 138-142: IEEE.

[23] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 19, no. 7, pp. 711-720, 1997.

[24] S. Nakariyakul, "High-dimensional hybrid feature selection using interaction information-guided search," *Knowledge-Based Systems*, vol. 145, pp. 59-66, 2018.

[25] A. Bhattacharjee *et al.*, "Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses," *Proceedings of the National Academy of Sciences*, vol. 98, no. 24, pp. 13790-13795, 2001.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

A Secure Mechanism to Protect UAV Communications

Seyed Aref Ayati, Hamid Reza Naji

Dept. of Computer Engineering and Information Technology, Graduate University of Advanced Technology, Kerman, Iran,
a.ayati@kgut.ac.ir, naji@kgut.ac.ir

***Abstract*— This paper presents a novel authentication method based on a distributed version of Kerberos for UAVs. One of the major problems of UAVs in recent years has been cyber-attacks which allow attackers to control the UAV or access its information.**
The growing use of UAVs has encouraged us to investigate the methods of their protection especially authentication of their users. In the past, the Kerberos system was rarely used for authentication in UAV systems. In our proposed method, based on a distributed version of Kerberos, we can authenticate multiple ground stations, users, and controllers for one or more UAVs. This method considers most of the security aspects to protect UAV systems mainly in the authentication phase and improves the security of UAVs and ground control stations and their communications considerably.

Index Terms — UAV, UAV Communication Network, Internet of Drones, Internet of Things, Smart Cities, IoT, Internet Security, Cryptography, Kerberos, Access Control, Distributed Systems

I. INTRODUCTION

An unmanned aerial vehicle (UAV) is an autonomous or remotely controlled (by Radio or Ground Control Station) vehicle without a driver or any crew in the sky [1].
In slang terms, UAV is also called a drone, but we should know that Drone has a more comprehensive meaning than UAV. For example, we can have a drone boat or drone ship or drone aircraft, etc. [2].
Now we should be noted in this article that the meaning of drones and UAVs is similar.

UAVs can be flown in two different ways. By using a remote-control system (RC) or using the ground control station (GCS) which has several crew on the ground. We have included both in a classification called ground station (GS) in a more comprehensive sense.

In an RC system, the user looks at the Drone or UAV and controls it or looks at the display that installs on the radio control of UAV or controls the UAV with a smartphone and its display. In these ways, users can control UAVs by mounted cameras on UAVs. In all models, we can try to be flown the UAV by real-time signals transmitted between user and UAV. For communication between UAV and user, we can use a lot of protocols such as telemetry, WIFI, cellular, and other networking protocols.

In the GCS model, the controller or crew uses computers and network protocols to connect the control software to the UAV and control it by cameras and sensors such as ultrasonic sensors. And UAVs and their crew have missions that they should pass it.
The sensors that are installed on UAV, help us to compute the location of UAV and we can compute its distance to us (with GPS sensor) and also helps us to compute the altitude of UAV (with ultrasonic sensors) and sensors can help us in other fields.
In the end, we can monitor the UAV by collecting this data [3].
In recent years, users use UAVs for scientific researches, posting boxes, agricultural use, making movies, technical use, and military use, and other uses. Also, users use UAVs to have smart cities in the Internet of Things era. We should know that the UAVs are classified by altitude range, weight, and flight endurance [4].
Example of UAV use: In film making industry, we can use UAVs to have Heli-shots video or in military use, we can use UAVs to give information from opponent military positions or

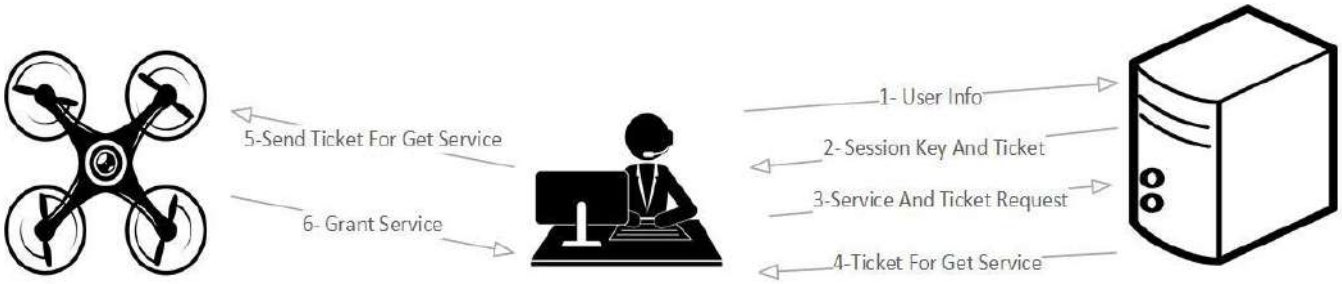


Figure 1. Simple Kerberos System

in the agriculture industry, we can use UAVs for field spraying, and also in other fields, we can use UAVs.

According to researches, in recent years the US military has increased its investment in drone research and production from $ 2.3 billion in 2008 to $ 4.2 billion by 2012 [5]. This proves that governments and countries are willing to invest in the development of this technology. However, one of the most important areas of this investment, which seems to be more important than other parts, is investing in their cyber security and paying attention to make them resistant against cyber-attacks.

In addition to using drones for civilian purposes (such as promotional videos or cargo or field spraying or even for personal and recreational use), drones for military purposes are also used in areas such as reconnaissance, border surveillance, logistics, and armed operations. This is one of the reasons why drones are so important to armies. According to this report, the security of drones in every army, every IoT field, and every smart city are well known. Most of the cyber-attacks on drones are aimed at obtaining information from their sensors, such as GPS sensors, or capturing images captured by them, or even attaining complete control of the drone.

II. GROUND CONTROL STATION (GCS)

GCS receives data from UAV and process these data, and at the end transforms the data from UAV and transfer it to other users in the same network via communication protocols [6]. Small unmanned aerial vehicles(sUAV) are controlled by portable devices such as Radio Control or mobile phone or laptops and other same devices and then operators or users can send information that gets from UAV to other users via network [7]. In recent years, the lack of secure and strong communication between the ground station and the UAV has been strongly felt, which is the basis of this article.

III. AUTHENTICATION USERS AND UAV

In the UAV area, we have UAV(s) and user(s) which need to access UAV(s) to control UAV(s) or get data from UAV(s). We like to have secure communication between UAV(s) and Authorized user(s). Because in the UAV area, we have a lot of types of attacks such as spoofing or modification, or forgery of the authorized user id. In 2011, the Iranian military attacked US Army RQ-170 UAV and could take the control of it [8]. So, in this paper, we propose a new solution which we hope that solve these problems.

1) Kerberos-Base Systems And Its Pros

An authentication protocol based on symmetric encryption and been used in a variety of systems proposed in the last years for use in network systems and distributed systems. In a distributed environment in which users at workstations wish to access services on servers distributed throughout the network, servers should be able to restrict access to authorized users and to be able to authenticate requests for service. In our method, the UAV plays the role of a destination server and the user plays the role of a client. This description can be seen in Figure 1.

2) Symmetric encryption method

In this method of encryption, simultaneously we have a key for the UAV and Kerberos-system.

This key is similar for UAV and Kerberos. These devices can encrypt a message with this unique key and then, can decrypt the message with this key. But we need transparency in this area.

Plain text: the original message text that we need to encrypt.

Encryption Algorithm: the algorithm for converting a plain text to ciphertext and that we show it at "E(c)" form.

Secret key: One of the inputs of the above algorithm is. The above algorithm uses it to convert plain text into ciphertext. This key is only for the receiver and the sender and no other agent should be aware of it.

Ciphertext: text taken from the output of the encryption algorithm. We call it "encrypted text" in social talks.

Decryption algorithm: its purpose is to produce the plain text of the ciphertext, which is, in a way, the inverse of the operation of the cryptographic algorithm.

To have symmetric cryptography, many algorithms have been proposed that most researchers, by the way, are studying and researching on their development and improvement. For example, algorithms such as DES, AES, RC4, IDEA, BLOWFISH, and so on, have been proposed, but when using them in UAVs, whether in terms of data encryption or authentication, we must keep in mind that Energy resources (such as batteries) in the UAV are limited, and also their processor speed must be considered (especially since we need to send information quickly with encryption and decryption operations to be done in the shortest time. that we should make decisions for the UAV and send it instantly the and then we receive the UAV response to our message instantly.)

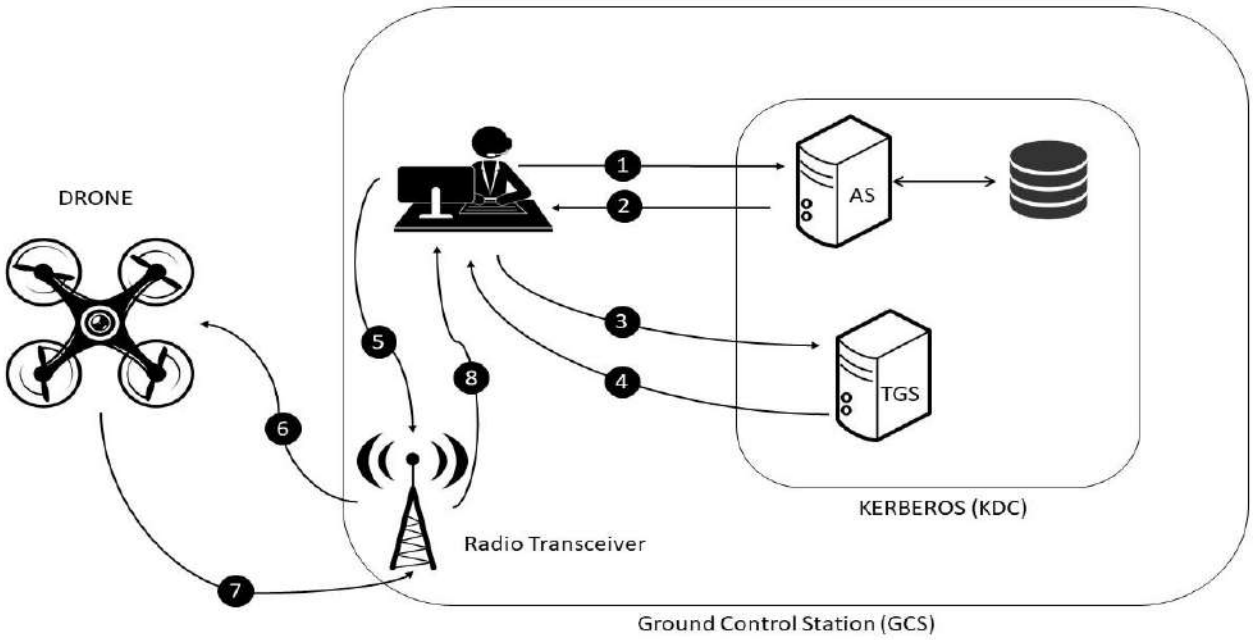


Figure 2. Our Proposed Method Scheme

In addition, other resources in the UAV also are limited, and moreover, if we want to add an authentication device to the UAV, we must keep in mind that the possibility of carrying weight for the UAV is very limited. As a result, we have to consider a completely economical and appropriate method of encryption and authentication. In this article, we present a method to cover the above issues.

IV. Related Works

In recent years, some researchers had research on the authentication field of UAV systems that we review some of them.

Zhang et al. proposed a lightweight AKA scheme (authentication and key agreement) in 2019 [9].

Li et al. proposed a lightweight communication mechanism that they believe is more secure and faster in comparison with the SM4 CTR algorithm. [10].

Pu et al. proposed a lightweight protocol for secure communication that uses a lightweight mac function for encryption based on physical unclonable function (PUF) [11].

Püllen et al. proposed a lightweight authentication scheme based on physical unclonable function and it especially proposed for vehicle network area. this scheme uses a symmetric encryption method that has needs low power and low computation [12].

Singh et al. proposed a lightweight authentication scheme for the IoT area that uses RSA asymmetric algorithm. but in the asymmetric encryption method, since we will use both public key and private key, we need to have a stronger CPU and have much more strong electrical power [13].

it is the downside of asymmetric cryptographic systems. Because we have limited electrical power resources in UAVs.

Base on research at [14] we have done a lot of research in recent years on the authentication field in the UAV area and researchers proposed any solution but some of their solutions have any problems that [14] talked about them.

Note that in the case of authentication, due to encryption, we consume a large amount of processor power, which means that we increase the battery consumption, which results in faster battery consumption, which results in reduced flight range of the UAV. In continuing, our idea is that if we have a system that transfers authentication out of the UAV, we will reduce processor consumption and it means we have tried to increase flight range.

In a nutshell, it seems that some of the protocols proposed in recent years have any vulnerabilities such as not complete security, have communication bit overhead problem, has password guessing attack risk, have information leakage, has session key violation, not having forward security, have poor performance, and server emulation and have other vulnerabilities. For example, in [15] we can see a comparison between common symmetric algorithms such as DES, 3DES, AES, and so on, that say these algorithms have any risk potential. And now, we like to say about an idea that proposes a novel method that can be a pattern in the next years for authentication in the UAVs area.

V. The Proposed Method

A. Kerberos-Base Authentication For UAV

we proposed our method, based on Figure 2. we will use a customize Kerberos system. In our proposed Kerberos systems, we can have any Ground Station (GS) or Ground Control Station (GCS) that we call “distributed systems”. This is one of the pros of using Kerberos systems for distributing elements.

In the Kerberos system, we can have one server (Authentication Server (AS)) to get a ticket for access to UAV, or we can have two servers to increase the security level that we call a second server, "ticket-granting server (TGS)". In the proposed method, we like to use two server models that we have both AS and TGS to get more security levels. As mentioned, for distributed systems, we can have multiple Kerberos bases, multiple users, and several UAVs. We use the Kerberos system because it gives us a higher level of security. That is, using a physical Kerberos device on the ground station (along with other devices we use to control the drone) and we will be using it to increase security and improve the quality of authentication and its security. The main feature of Kerberos is that in a distributed systems environment and where we have insecure communications, we can have an acceptable and secure authentication for authorized users to access the UAVs. Also, while restricting unauthorized users' access to the UAV, Kerberos provides a centralized authentication system that is both users authentication for UAV and UAV authentication for users. In insecure networks, one of the most common cyber-attacks that occur is authentication attacks, in which a third party, by changing its information to the information of an authorized user (we called it impersonation or spoofing authentication attack), intends to get access the UAV system, which in Kerberos system with Its countered. In the following, we will describe our proposed Kerberos system in three phases, and in the next section, we will clarify it.

TABLE I. Table Abbreviations Used in This Method

Abbreviations	Descriptions
ID_u	ID of user
ID_{UAV}	ID of UAV
AD_u	The network address of a user
U	User
UAV	UAV
K_{UAV}	Secret encryption key shared by AS & UAV
AS	Authentication Server
||	Concatenation
TS	Timestamp
Lifetime	the length of time for which the ticket is valid

1) Phase One (User Log on Session)

in step 1, the user needs the ticket for access to granting ticket server (TGS). So, the user sends a request for giving a ticket-granting ticket (TGT) from AS in message-1 format. In this message, the user sends his/her ID, password, and ticket-granting server (TGS) ID in the secure channel to get the ticket to access to ticket-granting server (TGS).
In step 2, based on the database information of AS, AS will respond in the secure channel to the user in message-2. If the user is authorized in As's database, he/she can get a ticket for access to TGS. In this message, we have the private key of the user and ticket to get access from TGS in an encrypted format.

2) Phase two (Type of Service)

In step 3, the user to get access to UAV, need to send to TGS, his/her information in message-3 format. In message-3, we have an ID of the user, ID of UAV, and ticket of TGS.
In step 4, TGS sends to the user, ticket for access to the UAV in message-4 format. In this message, we can see a ticket to get access to the UAV for the user.

3) Phase Three (Service Session)

In step 5, the user sends his/her ticket and ID in message-5 format to Ground Station Radio Transceiver and then, radio transceiver send's this information to UAV in the Wireless Sensor Network area in message-6 format. In these messages, we have an ID of the user and ticket to get access from UAV.
Notice: message-5 and message-6 are the same.
In step 6, if the user information that was sent to the drone via the radio receiver was correct, we will have the authorized access message (message -7) from the UAV for the user and it will be sent to the radio transceiver. And then, this information is sent to the user via radio transceiver in message-8 format.
In these messages, we have a notation that allows to access to UAV.

B. Clarification

In this method, we have authorized user(s) that need to access the UAV. For this get access, the user(s) needs to send a request for AS in the Kerberos system to get a ticket for sending it to TGS in the Kerberos system. In this step, the user(s) send a combination of the identifier of himself/herself, the identifier of TGS, and the password of himself/herself to AS. After sending a request from the user(s) to AS, AS will check the combination of an identifier of the user, identifier of TGS that user needs to access, and user password, and then if the user(s) is authorized (checks this info by information that saved in its database), AS sends a ticket to the user(s) that user(s) can access to TGS by this ticket. This is a first ticket (that type of this, is a ciphertext) that has a combination of an identifier of the user(s) and address of the user(s) and identifier of UAV that user(s) need to access to it and time stamp and lifetime with a secret key that both of users and UAVs will have it. In first, we should know just authorized users can use and reuse a ticket. And second, we should know this ticket, is encrypted by a private key and symmetric encryption method.
To continue, after the user(s) get a ticket for access to the TGS, the user(s) send to TGS a combination of an identifier of himself/herself, an identifier of UAV that they need access to, and a ticket of TGS that got from AS. TGS after getting the ticket of itself from the user(s), checks this ticket and if it's true, sends a new ticket to the user(s) that user(s) can use it by sending it to UAV and if UAV accepts it, the user(s) can get access to control of UAV. In this ticket (we call it second ticket

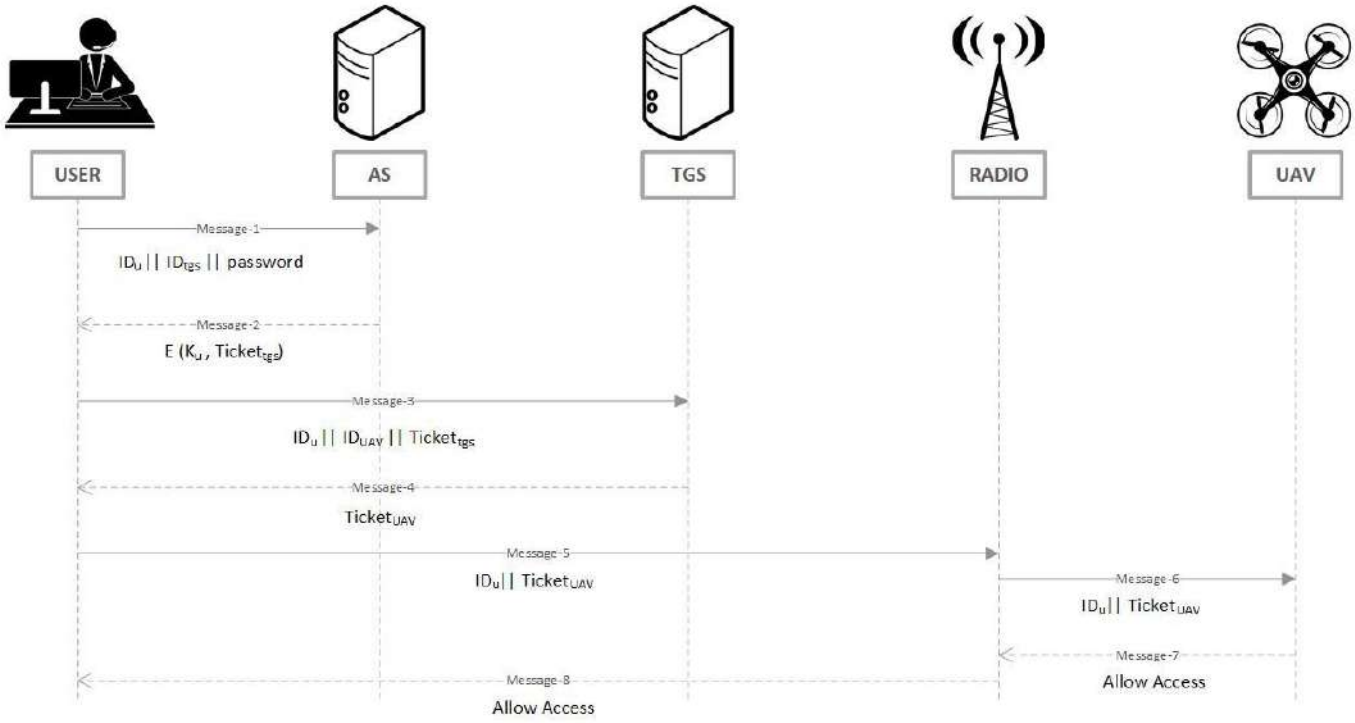


Figure 3. Sequence Diagram of Our Method

or UAV ticket which is a ciphertext) we have a combination of an identifier of the user, address of the user, identifier of UAV, timestamp-2, and lifetime-2 which is encrypted by the private key and symmetric encryption method. We have to keep in mind that the message, which includes the user ID and the TGS ticket, is sent to the drone by the ground radio stations, and in the meantime, there is no specific change in the data until it reaches the destination. In end, UAV gets access to the user(s) and all information collects by the UAV, will be sent to the user(s) and in addition, the user(s) can control the UAV or save data and information collected by UAV.

C. Parameters

We need to discuss the timestamp and lifetime that we used in both tickets. First of all, we need to know to counteract the forgery and spoofing of TGS information by an unauthorized third party (that third party can use the ticket of TGS by spoofing for reuse it), we will use timestamp and lifetime on our tickets [16].

Timestamp: Date and time of ticket issuance (example: 8.30 AM.).
Lifetime: Ticket life or validity and expiration time (example: five minutes.).
To continue, we have two definitions in which, we have two tickets that are encrypted with a secret key, and inside the encryption operation, we used a combination of data that we mentioned earlier. It should be noted that the symbol "E(c)" means the application of an encryption function on "C" data.

D. Definition-1

$Ticket_{tgs}$: E (K_{tgs} [ID_u || AD_u || ID_{tgs} || TS_1 || $LifeTime_1$]).

E. Definition-2

$Ticket_{UAV}$: E (K_{UAV} [ID_u || AD_u || ID_{UAV} || TS_2 || $LifeTime_2$]).

In the end, you can see all the steps of the mechanism of our proposed method in the diagram of Figure 3.
Note: if the user is not authenticated, then AS will not send any ticket toward the user(s).

VI. THE SECURITY ANALYSIS

In this method, we will use 4 rounds of Blowfish encryption, which is very fast in speed in operation and covers a flexible key length, which returns the encrypted ticket to the user. We use Blowfish because it seems faster than other symmetric key algorithms, has a low vulnerability, has flexibility, has low runtime versus more others, and has less memory usage based on [17], [18], and [19]. Previously said we need a low power consumption and with this mechanism, we had a Blowfish algorithm on Kerberos-system that has low power needed and our authentication server is on the ground. After using our proposed mechanism, because authentication location is on the ground, we have a UAV that has a processor with no need for encryption for authentication, and in the end, it means our UAV can fly more and more because it has a low power consumption. In addition, most attackers gain access to the UAV outside the ground station, but because we do most of the authentication operation on the ground station and the Kerberos-system is on the ground, there are fewer threats. Also in this method, since we use Timestamp and Lifetime, we experience a higher level

of security. Of course, the opponent would be able to reuse the ticket to spoof the TGS. But based on network security essential and standards book, to counter this vulnerability, the ticket includes a timestamp, indicating the date and time at which the ticket was issued, and a lifetime, indicating the length of time for which the ticket is valid (e.g., two hours) [16]. It should be noted again that in previous models, authentication was done by the UAV itself or used a simple control server, but in our proposed model, authentication will perform on the ground station by a Kerberos system, and we will experience more security. By comparison, as mentioned, most UAV authentication methods do not use the Kerberos system. They also use conventional encryption and authentication methods. We have proposed a Kerberos-based method, which we are further testing and will publish the results in future articles, but we anticipate to have good progress in terms of security compared to other methods.

VII. Conclusion and Future Works

In this paper, we proposed an authentication method based on a distributed version of Kerberos to protect UAV communications based on our researches. We believe that our proposed method improves the protection of UAVs considerably. Besides the ability to use symmetric cryptographic methods to increase the speed of confidentiality, we can use asymmetric ones to increase the level of confidentiality. Also, in new versions of Kerberos, with the addition of new security features, these systems can be greatly protected from intrusion. Other issues that we will pay attention to in the future are power consumption of both the ground station and the UAV, which we will try to cover by proposing lightweight encryption methods. In the future, by using this method, communication between several UAVs or several stations based on distribution systems can be established with more ease and security.

References

[1] A. C. Watts, V. G. Ambrosia, and E. A. Hinkley, "Unmanned Aircraft Systems in Remote Sensing and Scientific Research: Classification and Considerations of Use," Remote Sensing, vol. 4, no. 6, pp. 1671–1692, Jun. 2012.

[2] S. Herrick, "What's The Difference Between A Drone, UAV and UAS?,"www.botlink.com/blog/whats-the-difference-between-a drone-uav-and-uas, Nov. 2017.

[3] J. A. Marty, "Vulnerability analysis of the mavlink protocol for command and control of unmanned aircraft," Master's thesis, Air Force Institute of Technology, Air University, Wright-Patterson Air Force Base, Ohio, USA, 2013. [Online]. Available: https: //apps.dtic.mil/dtic/tr/fulltext/u2/a598977.pdf

[4] K. Dalamagkidis, "Classification of UAVs," in Handbook of Unmanned Aerial Vehicles, K. P. Valavanis, and G. J. Vachtsevanos, Eds. Dordrecht: Springer Netherlands, pp. 83–91, 2015.

[5] L. C. Baldor, " Flashy drone strikes raise status of remote pilots," Associated Press, https://www.bostonglobe.com/news/nation/2012/08/11/air-force-works-fill-need-for-drone-pilots/ScoF70NqiiOnv3bD3smSXI/story.html, Aug. 2012.

[6] N. AliKhan, N. Z. Jhanjhi, S. NawazBrohi, and A. Nayyar, "Emerging use of UAV's: secure communication protocol issues and challenges," in Drones in Smart-Cities. Amsterdam, The Netherlands: Elsevier, pp. 37–55, 2020.

[7] I. Bernik and B. Markelj, "Mobile devices and corporate data security," international Journal of Education and information Technologies, pp. 97-104, 2012.

[8] CNN wire staff, "Obama says U.S. has asked Iran to return drone aircraft", CNN, https://edition.cnn.com/2011/12/12/world/meast/iran-us-drone/index.html, 2011.

[9] Y. Zhang, D. He, L. Li, and B. Chen, "A lightweight authentication and key agreement scheme for Internet of Drones," Computer Communications, vol. 154, pp. 455–464, Mar. 2020.

[10] T. Li, J. Ma, X. Ma, C. Gao, H. Wang, C. Ma, J. Yu, D. Lu, and J. Zhang, "Lightweight Secure Communication Mechanism Towards UAV Networks," in 2019 IEEE Globecom Workshops (GC Wkshps), pp. 1–6, Dec. 2019.

[11] C. Pu and Y. Li, "Lightweight Authentication Protocol for Unmanned Aerial Vehicles Using Physical Unclonable Function and Chaotic System," in 2020 IEEE International Symposium on Local and Metropolitan Area Networks (LANMAN), pp. 1–6, Jul. 2020.

[12] D. Püllen, N. A. Anagnostopoulos, T. Arul, and S. Katzenbeisser, "Using Implicit Certification to Efficiently Establish Authenticated Group Keys for In-Vehicle Networks," in 2019 IEEE Vehicular Networking Conference (VNC), pp. 168–175, Dec. 2019.

[13] J. Singh, A. Gimekar, and S. Venkatesan, "An efficient lightweight authentication scheme for human-centered industrial Internet of Things," International Journal of Communication Systems, Nov. 2019.

[14] Y. Lei, L. Zeng, Y.-X. Li, M.-X. Wang, and H. Qin, "A Lightweight Authentication Protocol for UAV Networks Based on Security and Computational Resource Optimization," IEEE Access, vol. 9, pp. 53769–53785, 2021.

[15] S. Chandra, S. Paira, S. S. Alam, and G. Sanyal, "A comparative survey of Symmetric and Asymmetric Key Cryptography," in 2014 International Conference on Electronics, Communication and Computational Engineering (ICECCE), pp. 83–93, 2014.

[16] W. Stallings, "Network Security Essentials: Applications and Standards", 6th Edition. Pearson Education, 2016.

[17] M. Mitali, V. Kumar, and A. Sharma, "A Survey on Various Cryptography Techniques," International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), vol. 3, no. 4, pp. 307–312, 2014.

[18] O. G. Abood, S. K. Guirguis, "A Survey on Cryptography Algorithms," International Journal of Scientific and Research Publications (IJSRP), vol. 8, 2018.

[19] M. P. Chaudhari, S. R. Patel, "A Survey on Cryptography Algorithms," International Journal of Advance Research in Computer Science and Management Studies, vol. 2, 2014.

Seyed Aref Ayati received the B.S. degree of computer sciences from Isfahan University in 2020. He is currently a master student of Information Technology Engineering at the Graduate University of Advanced Technology, Iran. His research interests are computer networks and security, UAV systems, IT Management.

Hamid Reza Naji is associate professor of computer engineering at the Graduate University of Advanced Technology, Iran. His research interests include embedded systems, distributed, parallel and multi-agent systems, networks, and security. Dr. Naji has a Ph.D. in computer engineering from the University of Alabama in Huntsville, USA.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Adaptive TSK Fuzzy Terminal Sliding-Mode Control of Two Coupled Cart-Mounted Inverted Pendulums

1st Naser Padar, 2nd Mohammad Javad Mirzaei and 3rd Amir Abolfazl Suratgar

1st Department of Electrical Engineering, *Amirkabir University of Technology*, Tehran, Iran, padar@aut.ac.ir
2nd Faculty of Electrical and Computer Engineering), *University of Tabriz*, Tabriz, Iran, mjmirzaei95@ms.tabrizu.ac.ir
3rd Department of Electrical Engineering, *Amirkabir University of Technology*, Tehran, Iran, a-suratgar@aut.ac.ir

***Abstract*— Inverted pendulum and its variants, including two connected inverted pendulum on carts (TCIPC), are suitable benchmarks to study various control methods of underactuated systems. This work will use an adaptive fuzzy terminal sliding mode controller (AFTSMC) to investigate the stabilization problem of the TCIPC system. First, two finite-time sliding manifolds are proposed, and then a controller law is designed to drive the system states to the manifolds within a finite time, assuring the whole process to be finite-time. A Sugeno (TSK) type fuzzy controller is employed to address the chattering problem of the sliding mode controller. Simulation results show that the proposed AFTSMC method effectively stabilizes the error states and mitigates the chattering.**

***Index Terms*— terminal sliding mode control, two connected inverted pendulum on carts TCIPC, finite-time convergence, Sugeno TSK fuzzy tuner, chattering reduction**

I. Introduction

Most mechanical systems with classical rigid mechanisms are equal-actuated systems, which means that the number of actuators equals the degrees of freedom (DOF). The term underactuated refers to the systems which have fewer control inputs than the possible degrees of freedom. Underactuated systems are frequently encountered when a system has flexible mechanisms because not all degrees of freedom can be actuated and thus not directly controlled [1]. The control designer can also decide to omit one or more of the control inputs to simplify the controller design. Moreover, they can be found in case of a defective actuator [2]. Thus, underactuated control is vital for achieving fault-tolerant systems or reducing costs or complexity by using fewer control inputs. Examples of underactuated systems exist in a wide range of applications like underwater robots, UAVs [3] power systems, ball on the beam [4], and inverted pendulum (IP) [5]. Inverted pendulum and its variants, including single inverted pendulum (SIP), double inverted pendulum (DIP), rotary inverted pendulum (RIP), two-wheeled inverted pendulum (TWIP), cart-mounted inverted pendulum (IPC), and two connected cart-mounted inverted pendulums (TCIPC) are used to model several real-life devices [6], [7], [8]. Examples of such devices are two-wheeled robots [9] and overhead cranes [10]. Thus, studying control methods of inverted pendulum systems can give insight into how to control those devices. So far, researchers have investigated various methods for controlling inverted pendulums or their variants. Some examples include classical PID controllers [11], pole placement method [12], optimal and evolutionary algorithms [13], [14], feedback liearization [15], fractional order control [16], [17], and backstepping method [18]. Sliding mode control (SMC) has also been effectively employed to control inverted pendulum and various other nonlinear systems [19], [20], [21]. SMC has the following main characteristics:

- robustness: SMC is capable of rejecting disturbances and handling model uncertainties.
- fast response: various SMC methods exist in the literature that can guarantee finite-time stability.
- adjustable transient performance: by employing different surfaces and/or control laws, one can achieve desired transient performance.
- Simple implementation compared to some other control strategies.

Despite traditional SMC approaches having the same robustness capability, they can only guarantee an asymptotic convergence as they employ linear sliding surfaces [22]. Opposed to the traditional SMC, the terminal sliding mode control (TSMC) methods can achieve finite-time convergence and fast response due to special nonlinear surfaces in their design [23], [24], [25]. Thus, for this study, a TSMC method has been chosen as a basis. A second intention is to design the

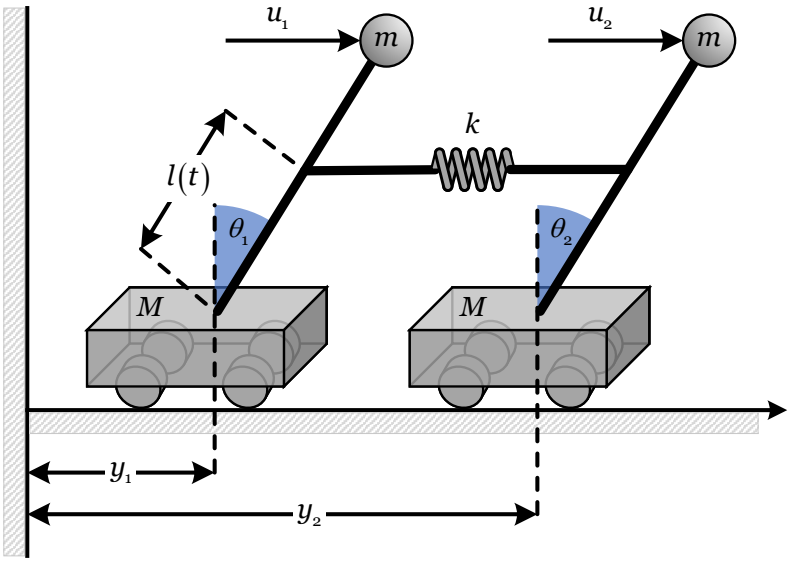


Fig. 1. Schematic of two connected inverted pendulums on carts

control law to ensure robustness against external disturbances and model uncertainties. While this will add robustness to the system, it will introduce chattering as a negative side effect. Discontinuous control input caused by the chattering phenomenon is a major limitation of robust SMC methods. The chattering problem seriously impacts the stability and reliability of the system. To overcome this issue, a fuzzy gain tuner is used to mitigate the chattering problem of sliding mode control. Additionally, the TCIPC system will serve as a benchmark to verify the proposed control method. Also, finite-time stability proof of the proposed adaptive fuzzy terminal sliding mode control (AFTSMC) method will be given. The rest of this paper is organized as follows. First, the model of TCIPC and required preliminaries are given in Section 2. In Section 3, the proposed AFTSMC method is described. The simulation results and discussions are detailed in Section 4, and finally, the conclusion is brought in Section 5.

II. Problem Formulation

The schematic of the TCIPC system is shown in Fig. 1. The pivot positions of the coupling spring (l) and the cart positions (y_1 and y_2) are assumed as functions of time. Here, the control objective is to bring and maintain the pendulums at the desired angle by applying the appropriate inputs u_1 and u_2. The following differential equations describe TCIPC dynamics [26]:

$$\begin{cases} \ddot{\theta}_1 = (\frac{g}{cL} - \varphi(t)l(t))\theta_1 - \frac{m}{M}\sin(\theta_1)\dot{\theta}_1^2 \\ \qquad + \varphi(t)l(t)\theta_2 + \varphi(t)(y_2 - y_1) + \frac{1}{cmL^2}u_1 \\ \ddot{\theta}_2 = (\frac{g}{cL} - \varphi(t)l(t))\theta_2 - \frac{m}{M}\sin(\theta_2)\dot{\theta}_2^2 \\ \qquad + \varphi(t)l(t)\theta_1 + \varphi(t)(y_1 - y_2) + \frac{1}{cmL^2}u_2 \end{cases} \tag{1}$$

$$c = \frac{m}{m+M}, \quad \varphi(t) = \frac{k(l(t) - cL)}{cmL^2}$$

In the equation above, θ_1 and θ_2 are the angles of inverted pendulums, $\dot{\theta}_1$ and $\dot{\theta}_2$ indicate angular velocities, g is the gravity constant, L is the length of the pendulums, m is mass of each pendulum, M is mass of each cart, and k is the spring constant. By considering $[x_1, x_2, x_3, x_4]^T = [\theta_1, \dot{\theta}_1, \theta_2, \dot{\theta}_2]^T$ the state-space model of the TCIPC is obtained as follows:

$$\begin{cases} \dot{x}_1(t) = x_2(t) \\ \dot{x}_2(t) = f_1 + \kappa_1 u_1(t) \\ \dot{x}_3(t) = x_4(t) \\ \dot{x}_4(t) = f_2 + \kappa_2 u_2(t) \end{cases} \tag{2}$$

where

$$\begin{aligned} f_1 =& (\frac{g}{cL} - \varphi(t)l(t))x_1 - \frac{m}{M}\sin(x_1)x_2^2 \\ & + \varphi(t)l(t)x_3 + \varphi(t)(y_2 - y_1) \\ f_2 =& (\frac{g}{cL} - \varphi(t)l(t))x_3 - \frac{m}{M}\sin(x_3)x_4^2 \\ & + \varphi(t)l(t)x_1 + \varphi(t)(y_1 - y_2) \end{aligned} \tag{3}$$

$$\kappa_1 = \kappa_2 = \frac{1}{cmL^2}$$

Additionally, l, y_1 and y_2 are assumed as in the following:

$$l(t) = sin(\omega t), \quad y_1 = sin(\omega_1 t), \quad y_2 = sin(\omega_2 t) + \Delta \tag{4}$$

Considering the tracking errors as $e_i = x_i - r_i$ $(i = 1, 2, 3, 4)$, dynamics of the error can be obtained as in the subsequent equation:

$$\begin{cases} \dot{e}_1 = e_2 \\ \dot{e}_2 = f_1 + \Delta E_1 + k(u_1 + d_1) \\ \dot{e}_3 = e_4 \\ \dot{e}_4 = f_2 + \Delta E_2 + k(u_2 + d_2) \end{cases} \tag{5}$$

where ΔE_i and d_i $(i \in \{1, 2\})$ are new terms added in spite of model uncertainties and disturbances, respectively. In the following, assumptions and lemmas which will be utilized in the controller design procedure have been presented.

Lemma 1. *An arbitrary non–autonomous dynamical system can be written as in the following equation:*

$$\dot{x} = f(x, t), \quad f : \mathbf{R}^n \to \mathbf{R}^n, \quad f(0, t) = 0 \tag{6}$$

Suppose $V(x) : \mathbf{R}^n \to \mathbf{R}$ to be a positive definite and proper Lyapunov-candidate-function. The system (6) is finite-time stable if the following condition holds:

$$\exists V(x) > 0 : \forall x \in X : \dot{V}(x) = \nabla V^T \cdot f(x, t) < -kV^a \tag{7}$$

where $k > 0$ and $a \in (0, 1)$. Hence, the system is finite-time stable and reaching time is bounded by T_{max} [27]:

$$T \leq T_{max} = \frac{V^{1-a}(0)}{k(1-a)} \tag{8}$$

Lemma 2. *Consider the system*

$$\begin{aligned} &\dot{x}_i = x_{i+1}, \quad i = 1 \cdots n-1 \\ &\dot{x}_n = -k_1 \operatorname{sig}(x_1)^{\alpha_1} - \cdots - k_n \operatorname{sig}(x_n)^{\alpha_n} \end{aligned} \tag{9}$$

where $\operatorname{sig}^a(x) = \operatorname{sgn}(x)|x|^a$ and $\operatorname{sgn}(x))$ denotes the sign function. $k_1, k_2, \cdots, k_n$ are positive numbers such that the polynomial $s_n + k_n s_{n-1} + \cdots + k_2 s + k_1 1$ is Hurwitz, and $\alpha_1, \cdots, \alpha_n$ satisfy

$$\begin{aligned} &\alpha_{i-1} = \frac{\alpha_i \alpha_{i+1}}{2\alpha_{i+1} - \alpha_i} \quad (i = 2, \cdots, n), \\ &\alpha_{n+1} = 1, \quad \alpha_n = \alpha \end{aligned} \tag{10}$$

There exists $\epsilon \in (0, 1)$ such that, for every $\alpha \in (1 - \epsilon, 1)$, the origin is a globally finite-time-stable equilibrium for the system (9).

Lemma 3. *For $\rho_i \in \mathbf{R}$, $i = 1, 2, ..., n$, $0 < r \leq 1$ the following inequality holds true [28].*

$$\Big(\sum_{i=1}^{n} |\rho_i| \Big)^r \leq \sum_{i=1}^{n} |\rho_i|^r \tag{11}$$

Assumption 1. *The uncertainties in the system dynamics are assumed to be bounded. Therefore, for the error system (5), we have:*

$$\exists (\sigma_1 > 0,\ \sigma_2 > 0) : (|\Delta E_1(t)| \leq \sigma_1) \wedge (|\Delta E_2(t)| \leq \sigma_2) \tag{12}$$

where σ_1 and σ_2 are known constants.

Assumption 2. *External disturbances are assumed to be bounded. Therefore, for the error system (5), we have:*

$$\exists (\varrho_1 > 0,\ \varrho_2 > 0) : (|d_1(t)| \leq \varrho_1) \wedge (|d_2(t)| \leq \varrho_2) \tag{13}$$

where ϱ_1 and ϱ_2 are known constants.

Model uncertainties and imposed disturbances should be satisfied by Assumptions (1) and (2) throughout the design procedure.

III. Control Design

This section discusses the design process of the proposed AFTSMC for finite-time tracking of the TCIPC system. SMC method generally consists of two steps. First, appropriate sliding surfaces are chosen to drive the states to the equilibrium point; then, a control law is designed to ensure that the system states reach the sliding surfaces and remain on them. The main objective of this study is to design a terminal sliding mode controller to drive error states to the equilibrium in the sense of finite-time, and the second objective is to mitigate the chattering by tuning controller gains using a Sugeno (TSK) fuzzy tuner.

A. Finite-time sliding manifolds

The following equations define the proposed finite-time integral sliding manifolds.

$$\begin{aligned} s_1 &= e_2 + \int_0^t \chi_1 d\tau, \quad \chi_1 = \alpha_1 \operatorname{sig}^{\gamma_1}(e_1) + \alpha_2 \operatorname{sig}^{\gamma_2}(e_2) \\ s_2 &= e_4 + \int_0^t \chi_2 d\tau, \quad \chi_2 = \alpha_3 \operatorname{sig}^{\gamma_3}(e_3) + \alpha_4 \operatorname{sig}^{\gamma_4}(e_4) \end{aligned} \tag{14}$$

Upon reaching error states to the above sliding manifolds, the conditions $s_1(t) = 0$ and $s_2(t) = 0$. Moreover, if the error states stay on the sliding surfaces, yields $\dot{s}_1(t)$, and $\dot{s}_2(t) = 0$. As a result, the error state dynamics can be described as:

$$\begin{cases} \dot{e}_1 = e_2 \\ \dot{e}_2 = -\alpha_1 \operatorname{sig}^{\gamma_1}(e_1) - \alpha_2 \operatorname{sig}^{\gamma_2}(e_2) \\ \dot{e}_3 = e_4 \\ \dot{e}_4 = -\alpha_3 \operatorname{sig}^{\gamma_3}(e_3) - \alpha_4 \operatorname{sig}^{\gamma_4}(e_4) \end{cases} \tag{15}$$

which, according to Lemma 2, consists of two decoupled finite-time stable subsystems.

B. Finite-time control law

Here the goal is to design u_1 and u_2 such that the error states (15) reach to the finite-time sliding manifolds (14) in a given finite time. The proposed control law is as follows:

$$u_i = -\kappa_i^{-1}(f_i + \chi_i + \eta_i \operatorname{sgn}(s_i)) \tag{16}$$

where the controller gain η_i is to be tuned by a Sugeno (TSK) type fuzzy tuner. The following section explains how the fuzzy logic controller (FLC) is implemented within the AFTSMC framework.

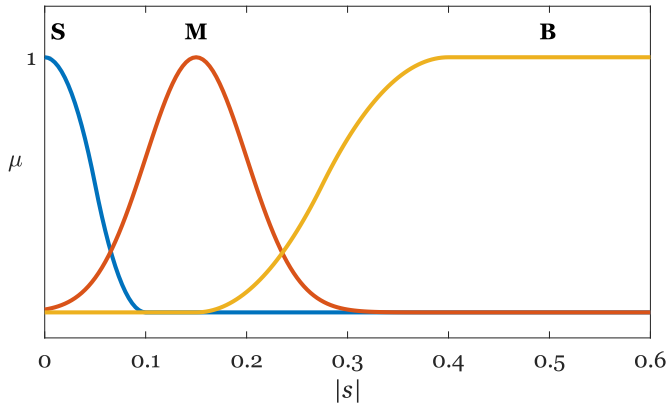


Fig. 2. Input membership function

C. TSK fuzzy tuner

The structure of a TSK type FLC is similar to that of a Mamdani type FLC; the only difference is in the consequence parts, where the output is a linear combination of the FLC inputs. To reduce the chattering, an FLC is employed to tune the controller gains η_1 and η_2, in which the absolute value of the sliding surfaces $s(t)$ is used as an input variable to the tuner. Fig. 2 illustrates the input fuzzy membership functions. In Fig. 2, the linguistic variables S, M, and B stand for small, medium, and Big, respectively. Based on experimental results, when $|s_i|$ is small, the controller gain can be reduced in order to reduce chattering. Therefore, TSK-type fuzzy tuner rules are defined as follows:

$$\begin{aligned} &\text{R1}: \quad \text{if } |s_i| \text{ is S} \quad \text{then} \quad \eta_i = m_1|s_i| + c_1 \\ &\text{R2}: \quad \text{if } |s_i| \text{ is M} \quad \text{then} \quad \eta_i = m_2|s_i| + c_2 \\ &\text{R3}: \quad \text{if } |s_i| \text{ is B} \quad \text{then} \quad \eta_i = m_3|s_i| + c_3 \end{aligned} \tag{17}$$

in which m_j and c_j $(j = 1, 2, 3)$ are positive constants. Generally, output of a TSK fuzzy controller can be interpreted as weighted average of the input variable:

$$\eta_i = \frac{\sum_{j=1}^{3} w_j(m_j|s_i| + c_j)}{\sum_{j=1}^{3} w_j} = \sum_{j=1}^{3} \bar{w}_j(m_j|s_i| + c_j) \tag{18}$$

where w_j is the firing strength of the j'th rule.

D. Stability analysis

In the proceeding section a stability proof based on Lyapunov finite-time stability theorem is carried out.

Theorem 1. *Consider the error dynamics (15), If the control inputs are chosen as (16), as a result the system states will reach sliding surfaces in the sense of finite-time.*

Proof: The Lyapunov function candidate is selected as following equation:

$$V = \frac{1}{2}\sum_{i=1}^{2} s_i^2 \tag{19}$$

Taking the derivative of V with respect to time, gives:

$$\dot{V} = \sum_{i=1}^{2} s_i\dot{s}_i = \sum_{i=1}^{2} s_i\big(f_i + \Delta E_i + \kappa_i(u_i + d_i) + \chi_i\big) \tag{20}$$

inserting u_i from (16) results in

$$\begin{aligned} \dot{V} &= \sum_{i=1}^{2} s_i\big(-\eta_i\,\mathrm{sgn}(s_i) + \Delta E_i + \kappa_i d_i\big) \\ &\le -\sum_{i=1}^{2}\big(\eta_i - |\Delta E_i| - \kappa_i|d_i|\big)|s_i| \end{aligned} \tag{21}$$

then by considering Equation (18) and Assumptions 1 and 2 the following inequality holds:

$$\dot{V} \le -\sum_{i=1}^{2}(\bar{c}_{\min} - \sigma_i - \kappa_i\varrho_i)|s_i| = -\sum_{i=1}^{2} p_i|s_i| \tag{22}$$

where $\bar{c}_{\min} = \min\{\bar{w}_1c_1, \bar{w}_2c_2, \bar{w}_3c_3\}$. By considering Lemma 3, the following inequality holds true

$$\dot{V} \le -\sum_{i=1}^{2} p_i\big(|s_i|^2\big)^{\frac{1}{2}} \le -p^*\Big(\sum_{i=1}^{2}|s_i|^2\Big)^{\frac{1}{2}} = -\sqrt{2}p^*V^{\frac{1}{2}} \tag{23}$$

where $p_i = \bar{c}_{\min} - \sigma_i - \kappa_i\varrho_i$ and $p^* = \min\{p_1, p_2\}$. Provided that $\bar{c}_{\min} > \sigma_i + \kappa_i\varrho_i$ according to the Lemma 1, using the proposed control (16), the stabilization problem of the error-state system (5) can be achieved within a finite time with T_r as its upper bound given as

$$T_r = \frac{\sqrt{2V_0}}{p^*}, \qquad V_0 = V|_{t=0} \tag{24}$$

Therefore, the proof is completed. ■

IV. Analysis and Numerical Simulation

In the current section, MATLAB/Simulink simulation results are given to demonstrate the effectiveness of the proposed method. To this end, the system parameters are selected as

$$\begin{aligned} &g = 9.8, \quad L = 1, \quad k = 1, \\ &M = 4, \quad m = 4, \quad \omega_1 = 2 \\ &\Delta = 2, \quad \omega_2 = 3, \quad \omega = 5 \end{aligned}$$

initial conditions and the reference values of the states are chosen as

$$\begin{aligned} x(0) &= [0 \quad 0.5 \quad 0 \quad -0.5]^T, \\ r \quad &= [\pi/8 \quad 0 \quad \pi/8 \quad 0 \quad]^T \end{aligned}$$

controller parameters are selected as follows

$$\begin{aligned} &\alpha_1 = 6, \quad \alpha_2 = 2, \quad \alpha_3 = 6, \quad \alpha_4 = 2 \\ &\gamma_1 = 19/39, \quad \gamma_2 = 19/29, \quad \gamma_3 = 19/39, \quad \gamma_4 = 19/29 \\ &m_1 = 50, \quad m_2 = 20, \quad m_3 = 10 \\ &c_1 = 0.3, \quad c_2 = 0.3, \quad c_3 = 0.3 \end{aligned}$$

imposed disturbances are considered as

$$d_1 = 0.2\sin(t), \quad d_2 = 0.2\sin(t)$$

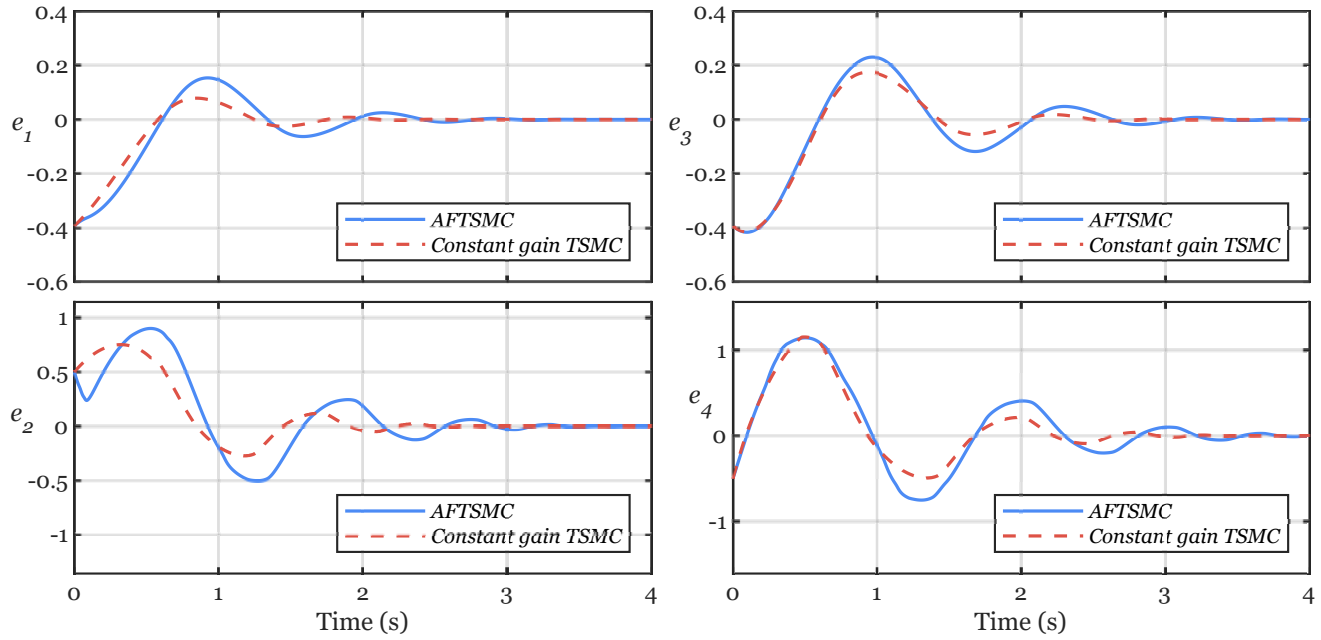


Fig. 3. Time response of the error states

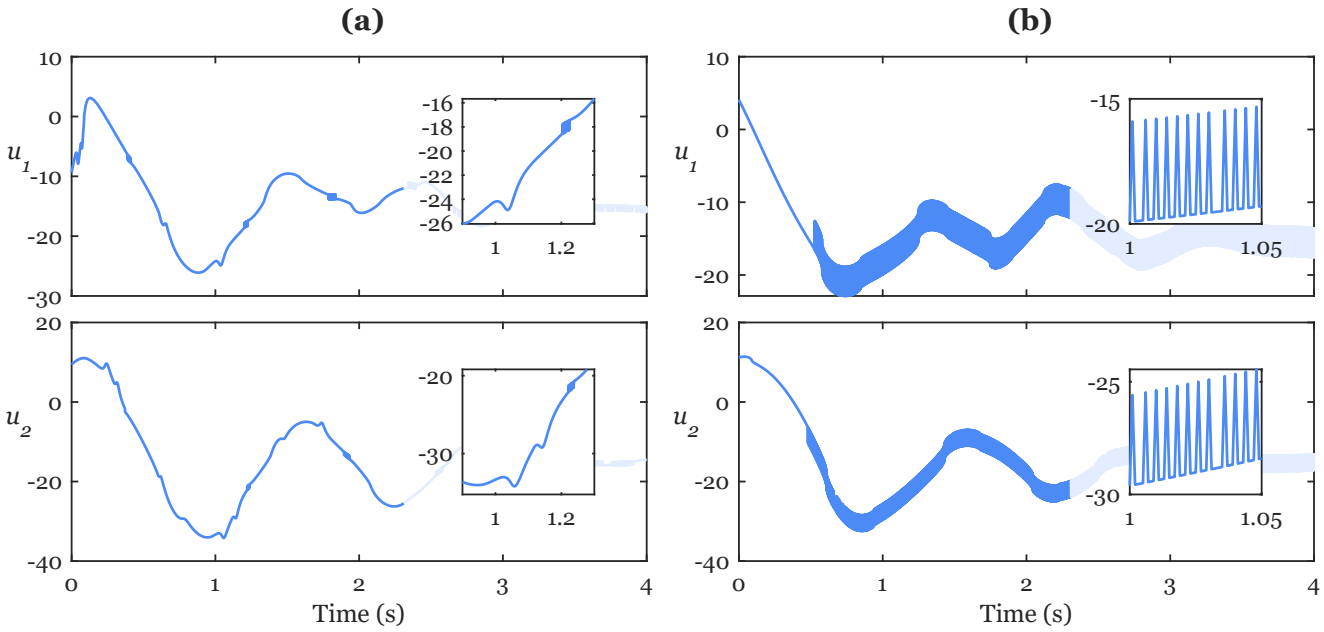


Fig. 4. Control Inputs (a) proposed AFTSMC method (b) TSMC with constant gain

Time responses of the error states are illustrated in Fig. 3. It can be seen that within 4 seconds, the error states converge to the origin. The performance of the proposed controller is also compared with the constant gain TSMC method in which the controller gains are $\eta_1 = \eta_2 = 1$. For a fair comparison, all other parameters are kept equal for both controllers. Fig. 4 compares control inputs with constant gain TSMC method. As can be seen, using the AFTSMC method, chattering in the control inputs has been effectively attenuated at the cost of slight changes in the settling times.

V. Conclusion

In this article stabilization problem for two connected inverted pendulums on carts was studied by the proposed AFTSMC method. The finite-time stability of the proposed method has also been provided. The designed control can significantly attenuate the chattering due to the TSK fuzzy controller acting as a gain scheduler. Simulation results confirmed the validity of the claims made about the proposed controller.

REFERENCES

[1] M. Valasek, "Design and control of under-actuated and over-actuated mechanical systems - Challenges of mechanics and mechatronics," *Vehicle System Dynamics*, vol. 40, no. SUPPL., pp. 37–50, 2004.

[2] M. Hutterer, D. Wimmer, and M. Schrodl, "Stabilization of a Magnetically Levitated Rotor in the Case of a Defective Radial Actuator," *IEEE/ASME Transactions on Mechatronics*, vol. 25, no. 6, pp. 2599–2609, 2020.

[3] M. Labbadi and M. Cherkaoui, "Robust adaptive nonsingular fast terminal sliding-mode tracking control for an uncertain quadrotor UAV subjected to disturbances," *ISA Transactions*, vol. 99, pp. 290–304, 2020.

[4] P. Ravichandran, S. Sathiyavathi, S. Sathish Babu, and A. Vimala Starbino, "Hybrid Arrangement of Iterative Learning Control Strategy for Ball and Beam System," *IETE Journal of Research*, 2020.

[5] W. Zhong and H. Röck, "Energy and passivity based control of the double inverted pendulum on a cart," in *IEEE Conference on Control Applications - Proceedings*, 2001, pp. 896–901.

[6] H. Gritli and S. Belghith, "Robust feedback control of the underactuated Inertia Wheel Inverted Pendulum under parametric uncertainties and subject to external disturbances: LMI formulation," *Journal of the Franklin Institute*, vol. 355, no. 18, pp. 9150–9191, 2018.

[7] A. Roose, S. Yahya, and H. Al-Rizzo, "Fuzzy-logic control of an inverted pendulum on a cart," *Computers and Electrical Engineering*, vol. 61, pp. 31–47, 2017.

[8] M. Hamza, H. Yap, I. Choudhury, A. Isa, A. Zimit, and T. Kumbasar, "Current development on using Rotary Inverted Pendulum as a benchmark for testing linear and nonlinear control algorithms," *Mechanical Systems and Signal Processing*, vol. 116, pp. 347–369, 2019.

[9] A. da Silva, Jr. and F. Sup, IV, "A Robotic Walker Based on a Two-Wheeled Inverted Pendulum," *Journal of Intelligent and Robotic Systems: Theory and Applications*, vol. 86, no. 1, pp. 17–34, 2017.

[10] H. Peng, B. Shi, X. Wang, and C. Li, "Interval estimation and optimization for motion trajectory of overhead crane under uncertainty," *Nonlinear Dynamics*, 2019.

[11] S. Valluru and M. Singh, "Stabilization of nonlinear inverted pendulum system using MOGA and APSO tuned nonlinear PID controller," *Cogent Engineering*, vol. 4, no. 1, 2017.

[12] M. Fauziyah, Z. Amalia, I. Siradjuddin, D. Dewatama, R. Wicaksono, and E. Yudaningtyas, "Linear quadratic regulator and pole placement for stabilizing a cart inverted pendulum system," *Bulletin of Electrical Engineering and Informatics*, vol. 9, no. 3, pp. 914–923, 2020.

[13] M. Gaya, A. Bisu, S. Salim, I. Madugu, L. Yusuf, M. Kaurangini, and M. Khairi, "Optimal PID design approaches for an inverted pendulum system," *International Review of Automatic Control*, vol. 9, no. 3, pp. 167–174, 2016.

[14] D. Al-Janan, H.-C. Chang, Y.-P. Chen, and T.-K. Liu, "Optimizing the double inverted pendulum's performance via the uniform neuro multiobjective genetic algorithm," *International Journal of Automation and Computing*, vol. 14, no. 6, pp. 686–695, 2017.

[15] H. Fukushima, M. Kakue, K. Kon, and F. Matsuno, "Transformation control to an inverted pendulum for a mobile robot with wheel-arms using partial linearization and polytopic model set," *IEEE Transactions on Robotics*, vol. 29, no. 3, pp. 774–783, 2013.

[16] R. Mondal, A. Chakraborty, J. Dey, and S. Halder, "Optimal fractional order PIλDμ controller for stabilization of cart-inverted pendulum system: Experimental results," *Asian Journal of Control*, vol. 22, no. 3, pp. 1345–1359, 2020.

[17] R. Mondal and J. Dey, "Performance Analysis and Implementation of Fractional Order 2-DOF Control on Cart-Inverted Pendulum System," *IEEE Transactions on Industry Applications*, vol. 56, no. 6, pp. 7055–7066, 2020.

[18] J. Huang, T. Zhang, Y. Fan, and J.-Q. Sun, "Control of Rotary Inverted Pendulum Using Model-Free Backstepping Technique," *IEEE Access*, vol. 7, pp. 96 965–96 973, 2019.

[19] E. Aslmostafa, M. Mirzaei, M. Asadollahi, M. Badamchizadeh, and S. Ghaemi, "Tracking of Two Connected Inverted Pendulum on Carts by Using A Fast Terminal Sliding Mode Control with Fixed-time Convergence," in *2021 7th International Conference on Control, Instrumentation and Automation, ICCIA 2021*, 2021.

[20] M. Mirzaei, E. Aslmostafa, M. Asadollahi, and M. Badamchizadeh, "Robust adaptive finite-time stabilization control for a class of nonlinear switched systems based on finite-time disturbance observer," *Journal of the Franklin Institute*, 2021.

[21] A. Nasir, R. Raja Ismail, and M. Ahmad, "Performance comparison between sliding mode control (SMC) and PD-PID controllers for a nonlinear inverted pendulum system," *World Academy of Science, Engineering and Technology*, vol. 46, pp. 400–405, 2010.

[22] M. Mirzaei, E. Aslmostafa, M. Asadollahi, M. Badamchizadeh, and S. Ghaemi, "Fuzzy Terminal Sliding Mode Control for Finite-time Synchronization of Electrostatic and Electromechanical Transducers," in *6th Iranian Conference on Signal Processing and Intelligent Systems, ICSPIS 2020*, 2020.

[23] A. Polyakov, D. Efimov, and W. Perruquetti, "Finite-time and fixed-time stabilization: Implicit Lyapunov function approach," *Automatica*, vol. 51, pp. 332–340, 2015.

[24] J. Mishra, X. Yu, and M. Jalili, "Arbitrary-Order Continuous Finite-Time Sliding Mode Controller for Fixed-Time Convergence," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 65, no. 12, pp. 1988–1992, 2018.

[25] H. Rabiee, M. Ataei, and M. Ekramian, "Continuous nonsingular terminal sliding mode control based on adaptive sliding mode disturbance observer for uncertain nonlinear systems," *Automatica*, vol. 109, 2019.

[26] Y.-J. Liu, W. Wang, S.-C. Tong, and Y.-S. Liu, "Robust adaptive tracking control for nonlinear systems based on bounds of fuzzy approximation parameters," *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans*, vol. 40, no. 1, pp. 170–184, 2010.

[27] S. Bhat and D. Bernstein, "Finite-time stability of continuous autonomous systems," *SIAM Journal on Control and Optimization*, vol. 38, no. 3, pp. 751–766, 2000.

[28] A.-M. Zou, K. Kumar, and A. de Ruiter, "Fixed-time attitude tracking control for rigid spacecraft," *Automatica*, vol. 113, 2020.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Adjoint relations of S-fuzzy posets with some categories

Leila Shahbaz

Department of Mathematics, *University of Maragheh*, *Maragheh, Iran,* leilashahbaz@yahoo.com, 1 shahbaz@maragheh.ac.ir

***Abstract*— In this paper, recalling the category FPos-S of all S-fuzzy posets, and action and order preserving maps between them, some categorical properties of the category FPos-S is studied. In particular, we characterize products, coproducts, equalizers, coequalizers, pullbacks and pushouts in this category. Also, all forgetful functors between the category FPos-S and the categories FPos of fuzzy posets, Pos-S, Pos, Act-S and Set are considered, and the existence of left and right adjoints of all mentioned forgetful functors are studied.**

***Index Terms*—fuzzy poset, S-fuzzy poset, free, cofree, adjoint relations**

I. Introduction

In the literature, many kinds of ordered algebras like pogroups, posemigroups, pomonoids, and etc. have been studied so far. Recently, Fakhruddin in [2, 3] has been studied the category of posets acted on by a pomonoid S, absolute flatness and amalgams of S-posets. After then many researchers have studied the properties of the category of S-posets. Historically, Zadeh, [5], generalized the concept of ordering as fuzzy ordering in the fuzzy framework, and since then, very researchers, were motivated to work on this concept and its applications to other areas of sciences. Recently, De Baets, Bodenhofer at al. studied fuzzy orders and investigated important results concerning them and their representations. In most of these studies the ordering relation is fuzzified, and the underlying set is crisp. Furthermore, T-norms, or more generally, residuated lattices are used by the authors. For example Boolean algebras, Heyting algebras, BL-algebras, and MV-algebras are residuated lattices. Motivated by the classical approach to S-posets, we investigate S-fuzzy posets by fuzzification of an ordering relation. The aim of this paper is to introduce the concept of the actions of a fuzzy ordered semigroup on a fuzzy ordered set (S-fuzzy poset) and study some properties of this category. particularly, we describe products, coproducts, equalizers, coequalizers, pullbacks and pushouts in this category. Also, we study all forgetful functors between the category **FPos**-S and the categories **FPos**, **Pos**-S, **Pos**, **Act**-S and **Set**, and their left and right adjoints are studied.

In rest of this section, we recall some of the notions about the category of S-acts and S-posets (see [1, 4]).

Consider S as a monoid with 1 as its identity. A set A equipped with an action $\lambda : A \times S \to A, (\lambda(a, s)$ is denoted by as) is called a (*right*) S-*act* if $a1 = a$ and $a(st) = (as)t, \forall s, t \in S$ and $a \in A$. A map $f : A \to B$ between S-acts is called an S-*map* if it is action preserving, that is for each $a \in A, s \in S$, $f(as) = f(a)s$. **Act**-S denotes the category of all S-acts and S-maps between them.

Recall that a monoid (semigroup) S which is also a poset where its partial order $\leq$ is compatible with the binary operation (that is, $s \leq t$, $s' \leq t'$ imply $ss' \leq tt'$) is called a *pomonoid* (*posemigroup*).

A poset A which is also an S-act with the action $\lambda : A \times S \to A$ which is order preserving, and $A \times S$ is considered as a poset with componentwise order is called a (*right*) S-*poset*. An S-*poset map* (or *morphism*) is an action and order preserving map between S-posets. Moreover, it is shown that regular monomorphisms are exactly *order embeddings*; (mono)morphisms $g : X \to Y$ such that $g(a) \leq g(a')$ if and only if $a \leq a'$, for all $a, a' \in X$. **Pos**-S denotes the category of all S-posets and S-poset maps between them.

Goguen introduced L-fuzzy set as a generalization of Zadeh's fuzzy set, with L being a complete residuated lattice.

A map $A : U \to L$ is called an L-fuzzy set (briefly, fuzzy set) A on U and L^U denotes all the L-fuzzy sets on U.

An L-fuzzy relation (briefly, fuzzy relation) R' on U is a map $R' : U \times U \to L$.
(1) R' is called *reflexive* if $R'(x, x) = 1$, $\forall x \in U$,
(2) R' is called *transitive* if $\bigvee_{y \in U} R'(x, y) * R'(y, z) \leq R'(x, z)$, $\forall x, z \in U$,
(3) R' is called *symmetric* if $R'(x, y) = R'(y, x)$, $\forall x, y \in U$,
(4) R' is called *antisymmetric* if $\forall x, y \in L, R'(x, y) = R'(y, x) = 1 \to x = y$.

Let X be a set and $e_X : X \times X \to L$ be a fuzzy relation. e_X is called a *fuzzy partial order* (briefly, *fuzzy order*) on X if it is reflexive, antisymmetric and transitive. The pair (X, e_X) is called a *fuzzy partially ordered set*, or briefly *fuzzy poset*. Let $(X', e_{X'}), (Y', e_{Y'})$ be two fuzzy posets. A map $g : (X', e_{X'}) \to (Y', e_{Y'})$ is called *fuzzy order preserving* if $e_{X'}(x, y) \leq e_{Y'}(g(x), g(y)), \forall x, y \in X'$. **FPos** denotes the category of all fuzzy posets with fuzzy order preserving maps between them.

Recall that if e_X is a fuzzy order on X then the inverse fuzzy relation $e^{-1} \in L^{X \times X}$ defined by $e^{-1}(x, y) = e(y, x)$ of e, $\forall x, y \in X$, is also a fuzzy partial order on X. Moreover, the symmetrization $e^s = e \wedge e^{-1}$ of e is a fuzzy partial order on X.

II. S-fuzzy posets

In this section, the category **FPos**-S of S-fuzzy posets are introduced and some non trivial examples of S-fuzzy posets are given.

Definition 1. *Suppose S is a semigroup. A fuzzy ordered semigroup is a semigroup $(S, .)$ with the fuzzy partial order $e_S : S \times S \to L$ such that for all $(s_1, s_2), (s_1', s_2') \in S \times S$, $e_S(s_1, s_2) \wedge e_S(s_1', s_2') \leq e_S(s_1 s_1', s_2 s_2')$.*

Definition 2. *Take $(S, ., e_S)$ as a fuzzy ordered semigroup and X as a fuzzy poset with the fuzzy order $e_X : X \times X \to L$ which is an S-act, too. Then X is called an S-fuzzy poset if*
(1) $e_S(s_1, s_2) \leq e_X(xs_1, xs_2)$ for all $(s_1, s_2) \in S \times S$ and $x \in X$;
(2) $e_X(x_1, x_2) \leq e_X(x_1 s, x_2 s)$ for all $(x_1, x_2) \in X \times X$ and $s \in S$.

Remark 1. *Let $L = \{0, 1\}$ be the truth values. Then the fuzzy orders on a set X are exactly the classical orders on X. Then the fuzzy posets on L are classical posets and so S-fuzzy posets are generalizations of S-posets. For a classical partial order $\leq$, a fuzzy order $e_X : X \times X \to L$ is given by*

$$e_X(x, y) = \begin{cases} 1 & \text{if } x \leq y \\ 0 & \text{otherwise.} \end{cases}$$

Definition 3. *Suppose (X, e_X) and (Y, e_Y) are two S-fuzzy posets. A map $f : (X, e_X) \to (Y, e_Y)$ is called S-fuzzy poset map if it is fuzzy order preserving ($e_X(x, y) \leq e_Y(f(x), f(y)), \forall x, y \in X$) and action preserving ($f(xs) = f(x)s, \forall x \in X$ and $s \in S$).*

FPos-S denotes the category of all S-fuzzy posets with S-fuzzy poset maps between them.

Definition 4. *Suppose (A, e_A) and (B, e_B) are two S-fuzzy posets. An action preserving map $g : (A, e_A) \to (B, e_B)$ is called*
(1) S-fuzzy poset order embedding if $e_A(x, y) = e_B(g(x), g(y)), \forall x, y \in A$.
(2) S-fuzzy poset order isomorphism if it is an onto fuzzy order embedding.

Example 1. *S-fuzzy poset order embedding is injective but, generally, the converse does not hold. Take $(\mathbf{1} \sqcup \mathbf{1} = \{0, 1\}, e_{\mathbf{1}\sqcup\mathbf{1}})$ where $e_{\mathbf{1}\sqcup\mathbf{1}}(0, 0) = e_{\mathbf{1}\sqcup\mathbf{1}}(1, 1) = 1, e_{\mathbf{1}\sqcup\mathbf{1}}(1, 0) = e_{\mathbf{1}\sqcup\mathbf{1}}(0, 1) = 0$, and $(\mathbf{2} = \{0, 1\}, e_{\mathbf{2}})$ where $e_{\mathbf{2}}(0, 0) = e_{\mathbf{2}}(1, 1) = 1 = e_{\mathbf{2}}(0, 1), e_{\mathbf{2}}(1, 0) = 0$ both taken as S-fuzzy posets over a one element fuzzy ordered pomonoid (S, e_S). Let $i : (\mathbf{1} \sqcup \mathbf{1} = \{0, 1\}, e_{\mathbf{1}\sqcup\mathbf{1}}) \to (\mathbf{2} = \{0, 1\}, e_{\mathbf{2}})$ be the identity S-fuzzy poset map. Obviously, i is injective but it is not an order embedding.*

III. Limits and colimits

In this section, some of the limits and colimits in the category of S-fuzzy posets are considered.

Limits. Suppose $\{(X_i, e_{X_i})_{i \in I}\}$ is a family of S-fuzzy posets. Define $e_X(x, y) = \bigwedge e_{X_i}(x_i, y_i)$ for $X = \prod_{i \in I} X_i$ and every $(x, y) = ((x_i)_{i \in I}, (y_i)_{i \in I}) \in \prod_{i \in I} X_i \times \prod_{i \in I} X_i$, then it is a fuzzy partial order on the product of $\{(X_i, e_{X_i})_{i \in I}\}$ and (X, e_X) becomes an S-fuzzy poset.

The terminal S-fuzzy poset is the singleton S-fuzzy poset, and the initial S-fuzzy poset is empty.

The equalizer of a pair $f, g : (X, e_X) \to (Y, e_Y)$ of S-fuzzy poset maps is given by $(E = \{x \in X | f(x) = g(x)\}, e_E)$ with the action and order inherited from (X, e_X).

The pullback of S-fuzzy poset maps $f : (X, e_X) \to (Z, e_Z)$ and $g : (Y, e_Y) \to (Z, e_Z)$ is the sub S-fuzzy poset $(P = \{(x, y) \in X \times Y | f(x) = g(y)\}, (e_X \wedge e_Y)|_P)$ of $(X \times Y, e_X \wedge e_Y)$ together with the restricted projection maps.

Colimits. S-fuzzy poset congruences play an essential role in studying the structure of S-fuzzy posets. Now, using S-fuzzy poset congruences, we construct some colimits of the category of S-fuzzy posets such as pushouts and coequalizers.

Let $\{(X_i, e_{X_i})_{i \in I}\}$ be a family of S-fuzzy posets. Define

$$e_X(x, y) = \begin{cases} e_{X_i}(x, y) & \text{if } x, y \in X_i \\ 0 & \text{otherwise} \end{cases}$$

for $X = \coprod_{i \in I} X_i$ and every $(x, y) \in \coprod_{i \in I} X_i \times \coprod_{i \in I} X_i$, then it is a fuzzy partial order on the coproduct of $\{(X_i)_{i \in I}\}$ and (X, e_X) becomes an S-fuzzy poset.

Let (X, e_X) and (Y, e_Y) be S-fuzzy posets, and $f, g : (X, e_X) \to (Y, e_Y)$ be S-fuzzy poset maps. The coequalizer of f and g is the quotient S-fuzzy poset $(Y/\nu(K), e_{\nu(K)})$ where $\nu(K)$ is an S-fuzzy poset congruence generated by the fuzzy relation K on (Y, e_Y) defined by

$$K(x, y) = \begin{cases} 1 & \text{if } x = y \text{ or } x = f(a), y = g(a) \text{ or } x = g(a), \\ & y = f(a) \\ 0 & \text{otherwise} \end{cases}$$

for all $a \in X$ and $x, y \in Y$.

Let $f : (X, e_X) \to (Y, e_Y)$ and $g : (X, e_X) \to (Z, e_Z)$ be S-fuzzy poset maps. Then the pushout of the pair f and g is the quotient S-fuzzy poset, $((Y \sqcup Z)/\nu(K), e_{\nu(K)})$, where $Y \sqcup Z$ is the coproduct of (Y, e_Y) and (Z, e_Z), and $\nu(K)$ is the S-fuzzy poset congruence generated by the following fuzzy relation on $Y \sqcup Z$,

$$K(y, z) = \begin{cases} 1 & \text{if } y = z \text{ or } y = (1, f(x)), z = (2, g(x)) \text{ or} \\ & y = (2, g(x)), z = (1, f(x)) \\ 0 & \text{otherwise} \end{cases}$$

for all $x \in X$ and $y, z \in Y \sqcup Z$.

IV. Adjoint relations

In this section, we study the existence of free and cofree objects in the categories **FPos**-S, **FPos**, **Pos**-S, **Pos**, **Act**-S and **Set**.

A. Adjoint relations of S-fuzzy posets with fuzzy posets

Definition 5. *Free S-fuzzy poset on a fuzzy poset (P, e_P) is an S-fuzzy poset (F, e_F) with a fuzzy order preserving map $\iota : (P, e_P) \to (F, e_F)$ such that the following universal property holds; given any S-fuzzy poset (A, e_A) and a fuzzy order preserving map $g : (P, e_P) \to (A, e_A)$ there exists a unique S-fuzzy poset map $\bar{g} : (F, e_F) \to (A, e_A)$ such that $\bar{g} \circ \iota = g$.*

Note that this concept is the generalization of the concept of a free S-poset on a poset, defined in [1]. Also one could define an adjoint pair between the categories **FPos**-S and **FPos**.

Theorem 1. *For a given fuzzy poset (P, e_P) and a fuzzy ordered monoid (S, e_S), $P \times S$, with the order $e_P \wedge e_S : (P \times S) \times (P \times S) \to L$ given by $e_P \wedge e_S((x_1, s_1), (x_2, s_2)) = e_P(x_1, x_2) \wedge e_S(s_1, s_2)$ and the action $(x, s)t = (x, st), \forall x_1, x_2, x \in P$ and $\forall s_1, s_2, s, t \in S$ is the free S-fuzzy poset on (P, e_P).*

Proof: With the action $(x, s)t = (x, st)$ and with the order $e_P \wedge e_S : (P \times S) \times (P \times S) \to L$ given by $e_P \wedge e_S((x_1, s_1), (x_2, s_2)) = e_P(x_1, x_2) \wedge e_S(s_1, s_2), \forall x_1, x_2, x \in P$ and $\forall s_1, s_2, s, t \in S$, $(P \times S, e_P \wedge e_S)$ becomes an S-fuzzy poset and the map $\iota : (P, e_P) \to (P \times S, e_P \wedge e_S)$ given by $\iota(x) = (x, 1)$ will be a universal fuzzy order preserving map. For, let $g : (P, e_P) \to (A, e_A)$ be any fuzzy order preserving map to an S-fuzzy poset (A, e_A). Then $\bar{g} : (P \times S, e_P \wedge e_S) \to (A, e_A)$ defined by $\bar{g}(x, s) = f(x)s$ is the unique S-fuzzy poset map with $\bar{g} \circ \iota = g$. ∎

Corollary 1. *The forgetful functor $U_1 : \mathbf{FPos} - S \to \mathbf{FPos}$ has a left adjoint $F_1 : \mathbf{FPos} \to \mathbf{FPos} - S$ (the free functor) given by $F_1(P, e_P) = (P \times S, e_P \wedge e_S)$.*

Definition 6. *An S-fuzzy poset (C, e_C) with a fuzzy order preserving map $\delta : (C, e_C) \to (P, e_P)$ is a cofree S-fuzzy poset on a fuzzy poset (P, e_P) if the following universal property holds; given any S-fuzzy poset (A, e_A) and a fuzzy order preserving map $\beta : (A, e_A) \to (P, e_P)$ there exists a unique S-fuzzy poset map $\bar{\beta} : (A, e_A) \to (C, e_C)$ with $\delta \circ \bar{\beta} = \beta$.*

Theorem 2. *Let (S, e_S) be a fuzzy ordered monoid and (P, e_P) be a fuzzy poset. Then the set $((P, e_P)^{(S,e_S)}, e)$, of all fuzzy order preserving maps from (S, e_S) to (P, e_P), with the order $e : (P, e_P)^{(S,e_S)} \times (P, e_P)^{(S,e_S)} \to L$ defined by $e(f_1, f_2) = \bigwedge_{s \in S} e_P(f_1(s), f_2(s))$ and the action $(f_1 s)(t) = f_1(st)$ for all $f_1, f_2 \in (P, e_P)^{(S,e_S)}$ and $s, t \in S$ is the cofree S-fuzzy poset on (P, e_P).*

Proof: One can check that with the order and the action defined above, $((P, e_P)^{(S,e_S)}, e)$ is an S-fuzzy poset. It is easily seen that the cofree map $\delta : (P, e_P)^{(S,e_S)} \to (P, e_P)$ given by $\delta(f) = f(1)$ is order preserving, since

$$\bigwedge_{s \in S} e_P(f(s), g(s)) = e(f, g) \leq e_P(\delta(f), \delta(g)) = e_P(f(1), g(1)).$$

Also, for a given fuzzy order preserving map $\beta : (A, e_A) \to (P, e_P)$ from an S-fuzzy poset (A, e_A), there exists a unique S-fuzzy poset map $\bar{\beta} : (A, e_A) \to (P, e_P)$ defined by $\bar{\beta}(a)(s) = \beta(as)$ such that $\delta \circ \bar{\beta} = \beta$. $\bar{\beta}$ is fuzzy order preserving, since $e_A(x, y) \leq e_A(xs, ys) \leq e_P(\beta(xs), \beta(ys)), \forall x, y \in A, s \in S$, and hence $e_A(x, y) \leq \bigwedge_{s \in S} e_P(\beta(xs), \beta(ys)) = e(\bar{\beta}(x), \bar{\beta}(y))$. ∎

Corollary 2. *The forgetful functor $U_1 : \mathbf{FPos} - S \to \mathbf{FPos}$ has a right adjoint $K_1 : \mathbf{FPos} \to \mathbf{FPos} - S$ (cofree functor) defined by $K_1(P, e_P) = ((P, e_P)^{(S,e_S)}, e)$.*

B. Adjoint relations of (S-)fuzzy posets with (S-)posets

Take $H : \mathbf{FPos} - S \to \mathbf{Pos} - S$ defined by $H(A, e_A) = (A, \leq)$, where $(A, \leq)$ is the underlying S-fuzzy poset of (A, e_A). Then H is a functor from $\mathbf{FPos} - S$ to $\mathbf{Pos} - S$. Also, define $K : \mathbf{Pos} - S \to \mathbf{FPos} - S$ given by $K(A, \leq) = (A, e_A)$, where

$$e_A(x, y) = \begin{cases} 1 & \text{if } x \leq y \\ 0 & \text{otherwise} \end{cases}$$

Then K is a functor from $\mathbf{Pos} - S$ to $\mathbf{FPos} - S$. It is shown that $K \dashv H$ and $H \dashv K$. Similar functors give an adjoint pair between the category of fuzzy posets and the category of posets.

C. Adjoint relations of S-fuzzy posets with S-acts

Theorem 3. *The forgetful functor $U_2 : \mathbf{FPos} - S \to \mathbf{Act} - S$ has a left adjoint $F_2 : \mathbf{Act} - S \to \mathbf{FPos} - S$ (free functor) given by $F_2(A) = (A, e_A)$, where $e_A : A \times A \to L$ is given by*

$$e_A(a, b) = \begin{cases} 1 & \text{if } a = b \\ 0 & \text{otherwise} \end{cases}$$

Remark 2. *There exists no right adjoint for the forgetful functor $U_2 : \mathbf{FPos} - S \to \mathbf{Act} - S$. Because, if $K' : \mathbf{Act} - S \longrightarrow \mathbf{FPos} - S$ is a right adjoint of U_2 then $H \circ K' : \mathbf{Act} - S \xrightarrow{K'} \mathbf{FPos} - S \xrightarrow{H} \mathbf{Pos} - S$ would be a right adjoint of the forgetful functor $U : \mathbf{Pos} - S \longrightarrow \mathbf{Act} - S$. But by the note after Theorem 17 of [1], there is not any right adjoint to the forgetful functor $U : \mathbf{Pos} - S \longrightarrow \mathbf{Act} - S$.*

D. Adjoint relations of S-fuzzy posets with sets

Definition 7. *The free fuzzy poset on a set X is a fuzzy poset (X, e_X), where $e_X : X \times X \to L$ is given by*

$$e_X(x, y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{otherwise} \end{cases}$$

together with the identity map $i : X \to (X, e_X)$ which satisfies the following universal property; given any fuzzy poset (P, e_P) and a map $g : X \to P$ there does exist a single fuzzy poset map $\bar{g} : (X, e_X) \to (P, e_P)$ such that $\bar{g} \circ i = g$.

Lemma 1. *The forgetful functor* $U' : \mathbf{FPos} \longrightarrow \mathbf{Set}$ *has a left adjoint* $F' : \mathbf{Set} \longrightarrow \mathbf{FPos}$ (*free functor*) *defined by* $F'(X) = (X, e_X)$*, where* $e_X : X \times X \to L$ *is given by*

$$e_X(x, y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{otherwise} \end{cases}$$

Theorem 4. *The forgetful functor* $U'_2 : \mathbf{FPos} - S \longrightarrow \mathbf{Set}$ *has a left adjoint* $F'_2 : \mathbf{Set} \longrightarrow \mathbf{FPos} - S$ *given by* $F'_2(X) = F_1 \circ F'(X) = F_1(X, e_X)$*. Precisely,* $(X \times S, e_X \wedge e_S)$ *with the fuzzy order* $e_X : X \times X \to L$ *given by*

$$e_X(x, y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{otherwise} \end{cases}$$

and and the action defined by $(x, s)t = (x, st), \forall x \in X, s, t \in S$ *is the free* S*-fuzzy poset on a set* X*.*

Remark 3. *There exists no right adjoint for the forgetful functor* $U'_2 : \mathbf{FPos} - S \longrightarrow \mathbf{Set}$*. For, let* $H' : \mathbf{Set} \longrightarrow \mathbf{FPos} - S$ *be a right adjoint for* U'_2*. Then* $H \circ H' : \mathbf{Set} \longrightarrow \mathbf{FPos} - S \longrightarrow \mathbf{Pos} - S$ *would be a right adjoint for a forgetful functor* $U : \mathbf{Pos} - S \longrightarrow \mathbf{Set}$ *which is a contradiction by Remark 16 of [1].*

REFERENCES

[1] Bulman-Fleming, S., Mahmoudi, M., *The category of S-posets*, Semigroup Forum **71**(3) (2005), 443-461.

[2] Fakhruddin, S.M., Absolute flatness and amalgams in pomonoids. Semigroup Forum, **33** (1986), 15-22.

[3] Fakhruddin, S. M., On the category of S-posets, Acta Sci. Math. (Szeged) **52** (1988), 85-92.

[4] Kilp, M., U. Knauer, A. Mikhalev, Monoids, Acts and Categories, Walter de Gruyter, Berlin, New York, 2000.

[5] Zadeh, L. A., Fuzzy sets, Inform. Control, **8** (1965), 338-353.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

An Aggregated Revenue-Driven VPP Model Based on Marginal Price Tracking for Profit Maximization

Zohreh Azani[1], Hossein Shahinzadeh[1], Shohreh Azani[1], Gevork B. Gharehpetian[1], Ersan Kabalci[2], Mohamed Benbouzid[3]

[1] *Department of Electrical Engineering, Amirkabir University of Technology, Tehran, Iran,*
[2] *Department of Electrical Engineering, Nevsehir Haci Bektas Veli University, Nevsehir, Turkey.*
[3] *Institut de Recherche Dupuy de Lôme, University of Brest, 29238 Brest, France.*
z.azani@ieee.org, h.s.shahinzadeh@ieee.org, sh.azani@ieee.org, grptian@aut.ac.ir, kabalci@nevsehir.edu.tr, Mohamed.Benbouzid@univ-brest.fr

***Abstract*— Today, the benefits of using distributed generation (DG) sources at different voltage levels in smart grids cannot be ignored. DGs play a vital role in increasing the flexibility of operation in electricity grids. In addition, Dispatchable DGs can help to efficiently compensate for the imbalance caused by the uncertainty of renewable sources. On the other hand, renewable DGs can help to mitigate the final electricity price considering the impacts of emission. The benefits presented in most studies by researchers focus on the independent use of these resources. The negative facets associated with distributed generation sources, such as small-scale and intermittent generation capacity, cause hurdles in participating in the electricity market. This study aims to provide a solution to deal with these problems. In addition, a new energy management system is introduced and actions that can maximize the profit of the virtual power plant are evaluated. Then, the objective function and constraints related to maximizing the profit of the virtual power plant are introduced, and the proposed model is simulated using the Gravitational Search Optimization Algorithm (GSA) in MATLAB software. Finally, the simulation results for the 18-bus IEEE test distribution system are analyzed and the advantage of integrating distributed generation resources with the virtual power plant scheme is determined and discussed.**

***Index Terms* —Virtual Power Plant, Bilateral Contract, Maximum Profit, Distributed Generation Resources, Energy Management System.**

Nomenclature

P_t^D (MW)	Active power demand of VPP clients at each interval
$\lambda_t^{DSO.charge}$ $(\$/MWh)$	Electricity price for VPP clients at each interval
BC_t (MW)	Active power delivered to bilateral contract subscribers at each interval
λ_t^{BC} $(\$/MWh)$	Energy prices in the bilateral contract for each time period
$P_{kt}^{Upstream}$ (MW)	Total VPP power exchange with the upstream electricity market through grid supply point
λ_{kt}^{LMP} $(\$/MWh)$	Energy price at bus k
P_{it}^{DG} (MW)	Active power generation of the i^{th} DG at period t
$\lambda_i^{DG.cost}$ $(\$/MWh)$	Cost of active power generation of the i^{th} DG at period t
$y_{it}^{DG.start}$	Start-up indicator for the i^{th} DG at period t
$\lambda_i^{DG.startcost}$ $(\$)$	Start-up cost of the i^{th} DG
$Z_{it}^{DG.shut}$	Shutdown indicator for the i^{th} DG at period t
$\lambda_i^{DG.shutcost}$ $(\$)$	Shutdown cost of the i^{th} DG
P_{jt}^{USG} (MW)	Uncertain stochastic generation (USG) of uncertain unit j at interval of t
$\lambda_j^{USG.cost}$ $(\$/MWh)$	Generation cost of uncertain unit for stochastic generation at period t
P_t^{FL} (MW)	The amount of load shedding imposed on flexible loads at each interval
$\lambda_t^{FL.cost}$ $(\$/MWh)$	The cost of reducing the demand for flexible loads per interval
$P^{D.Max}$ (MW)	Maximum amount of power requested by VPP subscribers
$P^{BC.Max}$ (MW)	Maximum demanded power of subscribers, resulting from a bilateral contract
$P_k^{SS.max}$ (MW)	Maximum power capacity exchanged through the bus k with the upstream grid
$P_i^{DG.Max}$ (MW)	Maximum generation capability of the i^{th} DG
$P_j^{USG.Max}$ (MW)	Maximum uncertain generation capability of the j^{th} USG
$P^{FL.Max}$ (MW)	Maximum load shedding level
v_{it}^{DG}	Binary state variable of the i^{th} DG (1 if the DG unit is on,0 if the DG unit is off)
v_{jt}^{USG}	Binary state variable of the j^{th} USG (1 if the DG unit is on,0 if the DG unit is off)
rur_i	Ramp up rate limit of the i^{th} dispatchable DG
rdr_i	Ramp down rate limit of the i^{th} dispatchable DG
UT_i^{DG} (h)	Minimum up time of the i^{th} DG

DT_i^{DG} (*h*)	Minimum down time of the i^{th} DG
$E_t^{contract}$ (*MWh*)	The amount of energy traded according to bilateral contract

I. INTRODUCTION

In the last two decades, much attention has been paid to the integration of distributed energy resources (DERs). As can be seen now, most countries have strategies for investing in renewable energy (RES) based on recent forecasts for fossil fuels, global warming, and rising fossil fuel prices [1]. Because the return on investment in this part of the generation is slower than in power plants that use fossil fuels, governments are implementing various incentive schemes[2]. Despite the increasing role of distributed generation resources and various advantages over conventional power plants, problems have also been observed for the exploitation of these resources in the power system [3-5]. The main problems, which are due to the intermittent generation and less generation capacity, make these resources seem invisible to conventional power plants, which creates technical and commercial problems. Technical problems are reflected in the dynamic behavior of the power system, and commercial problems are also seen in the participation in the electricity market [5-6]. Today, one of the challenges in the electricity industry is how to control and operate DERs with the problems raised; therefore, it is necessary to consider improving system performance as well as uncertainties [7]. The development of new ways to manage generation and create the right infrastructure to participate in the electricity market is essential. Various methods have been proposed to deal with these problems, such as microgrids schemes and virtual power plants (VPP) [8-9]. The collection of distributed generation resources to provide storage capacity is a suitable solution to compensate for unexpected RES fluctuations [10]. Many researchers suggest a virtual power plant that includes a large number of DERs, so that the potential is equivalent to conventional power plants, so this organization overcomes the problems of invisibility and unreliability and becomes a sustainable economy [11]. Until now, DERs participated in the market individually, but after the emergence of VPP, independent DERs sold their generation to the organization by providing an estimate of their generation [12-13]. In this way, independent DERs participate in the market as a cooperative. VPPs are divided into two types: Technical Virtual Power Plant (TVPP) and Commercial Virtual Power Plant (CVPP) according to their purpose. TVPPs address how power generation and transmission, environmental issues, and technical problems with aggregation of DERs and CVPPs on how to plan and participate in the market and goals such as maximizing profits or minimizing generation costs with a variety of methods and models [14-16].

VPPs can participate in the market as a single power plant to manage part of the distribution network and aggregating resources, including conventional small power plants, distributed generation resources, storage facilities, consumers, and flexible loads [17]. For this purpose, VPPs need to build a supervisory station called the Energy Management System (EMS) to employ trained operators to manage the entire system according to their goals. In the following, the performance of VPP for the distribution system is specified and then the target plan and constraints to achieve maximum profit are presented, and finally, the results of the simulation are discussed [18-20].

II. OBJECTIVE FUNCTION

The VPP profit maximization model is formulated as Eqs. (1). In the objective function used in Eq. (1), the total benefit or expected revenue from the total set minus operating, and maintenance costs for a given period is calculated. The first part represents the total income from the sale of energy to ordinary and contract customers, and the second part is the income from the sale or purchase cost of energy at market prices from different networks of the national grid. The third section shows the cost of generating, start-up, and shutdown of DGs. The fourth part shows the cost of installation, operational, and maintenance of stochastic units, and the last part shows the cost of limiting flexible loads or loads shedding. In nomenclature, the concepts and equivalents of the symbols and abbreviations used in the objective function and constraints are shown.

$$MaxProfit = \left\{ \begin{array}{l} \sum_{t=1}^{T} (P_t^D . \lambda_t^{DSO.charge} + BC_t . \lambda_t^{BC} \\ + \sum_{K \in GSP} P_{kt}^{Upstream} . \lambda_{kt}^{LMP} - \sum_{i \in DG} (P_{it}^{DG} . \lambda_i^{DG.cost} \\ + y_{it}^{DG.start} . \lambda_i^{DG.start\,cost} + Z_{it}^{DG.shut} * \lambda_i^{DG.shut\,cost}) \\ - \sum_{j \in SG} P_{jt}^{USG} * \lambda_j^{USG.cost} - P_t^{FL} * \lambda_t^{FL.cost}) \end{array} \right\} \quad (1)$$

It should also be noted that when a VPP does not have accurate planning or accurate forecasting of uncertain parameters, it may face generation shortages; in this case, market wholesalers are one of the variables to help VPP to feed the load of subscribers.

A. Constraints for profit maximization model

The amount of power that a VPP exchange with the national grid per hour is limited by the amount of power requested and the maximum capacity of the feeders connected to the national grid. Eqs. (2) and (3) model the capacity of feeders or the capacity limitation of transformers or networks.

$$\left\{ \begin{array}{l} P_{kt}^{Upstream} \geq -min\{max\{((P^{D.\max} + P^{BC.\max}) \\ -\sum_{i \in DG} P_i^{DG.\min} - \sum_{j \in SG} P_j^{USG.\min} - P^{FL.\min})\}.P_k^{SS\max}\} \end{array} \right. \quad (2)$$

$$\left\{ \begin{array}{l} P_{kt}^{Upstream} \geq min\{max\{(\sum_{i \in DG} P_i^{DG.\max} + \sum_{j \in SG} P_j^{USG.\max} \\ +P^{FL.\max} - P^{D.\min} - P^{BC.\min})\}.P_k^{SS\max}\} \end{array} \right. \quad (3)$$

Eq. (2) is activated when the VPP wants to buy power from the network; Eq. (3) is also activated when the VPP sells power to any part of the national grid. Eq. (4) indicates that the amount of power of the subscribers under the VPP must be provided.

$$\sum_{i \in DG} P_{it}^{DG} + \sum_{j \in SG} P_{jt}^{USG} + P_t^{FL} + \sum_{K \in GSP} P_{kt}^{Upstream} \geq P_t^D + BC_t \quad \forall t \quad (4)$$

Eqs. (5) and (6) show the range of change of flexible loads in each period; which this share is considered as demand reduction or load shedding [21].

$$P_t^{FL} \leq P_t^{FL.\max} \qquad \forall t \quad (5)$$

$$P_t^{FL} \geq 0 \qquad \forall t \quad (6)$$

Eqs. (7) and (8) express the power range of the four dispatchable generation units located in P [22].

$$P_{it}^{DG} \leq P_i^{DG.\max} . v_{it}^{DG} \qquad \forall i.t \quad (7)$$

$$P_{it}^{DG} \geq P_i^{DG.\max}.v_{it}^{DG} \qquad \forall i.t \qquad (8)$$

Eqs. (9) and (10) indicate the capacity of uncertain stochastic units or renewables. In P_{jt}^{USG}, it is assumed to be a stochastic parameter and, as mentioned, may be predicted by DG owners [23].

$$P_{jt}^{USG} \leq P_j^{USG.\max}.v_{jt}^{USG} \qquad \forall j.t \qquad (9)$$

$$P_{jt}^{SG} \geq 0 \qquad \forall j.t \qquad (10)$$

Eqs. (11) and (12) show the slope of the output power changes of four dispatchable units [24].

$$P_{it}^{DG} - P_{i(t-1)}^{DG} \leq rur_i \qquad \forall i.t \qquad (11)$$

$$P_{it}^{DG} - P_{i(t-1)}^{DG} \leq -rdr_i \qquad \forall i.t \qquad (12)$$

Eqs. (13) and (14) of the model show them on and off status and prevent simultaneous entry and exit of dispatchable units.

$$y_{it}^{DG.start} - z_{it}^{DG.shut} = v_{it}^{DG} - v_{i(t-1)}^{DG} \qquad \forall i.t \qquad (13)$$

$$y_{it}^{DG.start} + z_{it}^{DG.shut} \leq 1 \qquad \forall i.t \qquad (14)$$

As mentioned earlier, bad predictions triggered by stochastic units cause VPPs to receive detriments. Hence, subject to tackle this problem, the VPP operator allocates a certain share of dispatchable units on a small scale as a security margin, which is given in Eq. (15).As a result, as the number of renewable units increases, the need for storage capacity will be proportionally increased.

$$\begin{cases} \sum_{i \in DG} P_i^{DG.\max} - P_{it}^{DG} + P_t^{FL} \geq \\ \zeta_0.(\sum_{j \in SG} P_{jt}^{USG}) + \zeta_{total}.(P_t^{FL} + \sum_{i \in DG} P_{it}^{DG}) \end{cases} \qquad \forall i.t \qquad (15)$$

ξ_0 is a percentage of the additional reserve due to bad forecasts to compensate for losses due to stochastic units, which is considered equal to 10%, And ξ_{total} is a small percentage of the total generation of dispatchable units and flexible loads, which is considered to be 5%.

Other characteristics for dispatchable units are Eqs. (16) and (17), which are shown in the minimum up and downtime constraints, respectively [25].

$$\begin{cases} v_{it}^{DG} - v_{i.(t-1)}^{DG} \leq v_{i.(t+TU_{iw})}^{DG} \\ .TU_{iw} = \begin{cases} w & w \langle UT_i^{DG} \\ 0 & w \rangle UT_i^{DG} \end{cases} \end{cases} \qquad \forall t.\forall i \qquad (16)$$

$$\begin{cases} v_{i.t(t-1)}^{DG} - v_{it}^{DG} + v_{i.(t+TD_{iw})}^{DG} \leq 1 \\ .TD_{iw} = \begin{cases} w & w \langle DT_i^{DG} \\ w & w \rangle DT_i^{DG} \end{cases} \end{cases} \qquad \forall t.\forall i \qquad (17)$$

Eqs. (18) and (19) show the power changes resulting from bilateral contracts. Δ_{BC} represents the allowable difference between the energy delivered for the bilateral contract, which is equal to 5% agreed, which this amount of deviation in the amount of delivered power is acceptable. It should also be noted that the amount of power delivered in 24 hours is equal to the amount of power concluded in the bilateral contract, which is shown in Eq. (20). It is also assumed that there is a linear relationship between market price and normal subscriber demand shown in Eq. (21).

$$BC_t \leq (1+\Delta_{BC}).E_t^{contract} \qquad (18)$$

$$BC_t \geq (1-\Delta_{BC}).E_t^{contract} \qquad (19)$$

$$\sum_{t=1}^{24} BC_t = \sum_{t=1}^{24} E_t^{contract} \qquad (20)$$

$$P_t^{Demand} = a.\lambda_t^{DA.forecast} + b(ie.a = .07, b = 8) \qquad (21)$$

III. Gravitational Search Algorithm

Gravity Search Algorithm (*GSA*) is an algorithm that depends on cooperative intelligence, but without memory. This optimization algorithm is designed with the aid of gravity laws and the motion of factors in an artificial system that is in discrete time. In the algorithm, the system's space is defined as the same as the problem. Based on the gravity law, every factor (mass) understands the situation and status of the other masses with the gravity law. Therefore, this force can be used as a tool for exchanging information. For example, all the forces exerted on mass M_1, by the other masses, are shown in Fig. 1 [26].

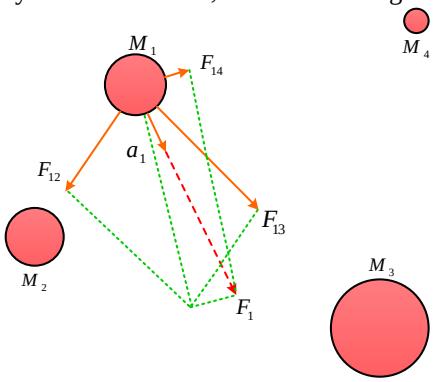


Figure 1. Forces exerted on mass M_1

In this algorithm, the mass of factors is determined by considering the target function. In a system with *n* masses, the situation of every mass is a point in the space. The situation of the i^{th} mass is shown by X_i.

$$X_i = \left(X_i^1, \ldots, X_i^d, \ldots, X_i^n\right) \qquad (22)$$

where *n* is the dimension of the problem and X_i^d is the dimension *d* of the i^{th} mass. In this system, the force $F_{ij}^d(t)$ is exerted on the mass *i*, by the mass *j*, in direction *d*, at time *t*. The amount of the force is calculated by Eq. (23).

$$F_{ij}^d(t) = G(t)\frac{M_{pi}(t) \times M_{aj}(t)}{R_{ij}(t) + \varepsilon}\left(x_j^d(t) - x_i^d(t)\right) \qquad (23)$$

In the above equation: $M_{aj}(t)$ is the active gravity mass of the mass *j*, $M_{pi}(t)$ is the passive gravity mass of the mass *i*, $G(t)$ is the constant gravity at the time *t*, R_{ij} is the distance between the masses *i* and *j*, ε is a very small number.

The amount of $M_{aj}(t)$ and $M_{pi}(t)$ are considered the same and equal to *M* in the algorithm. The gravitational constant gets a much more appropriate parameter for control of exploration and exploitation abilities in this algorithm, which is expressed as Eq. (24):

$$G(t) = G_0\, e^{\frac{-\alpha t}{T}} \qquad (24)$$

In which: G_0 and α are controlling coefficients, *T* is the system's life length, i.e., the whole number of the algorithm's repetition. The force which is exerted on the mass *i*, in the direction *d*, at the time *t*, based on the Eq. (25), equals the sum of the individual forces exerted by the other masses on it.

$$F_i^d(t) = \sum_{j=1, j \neq i}^{N} rand_j F_{ij}^d(t) \qquad (25)$$

In this equation, $Rand_j$ is a random number with the same distribution in the range of (1,0) that is considered to preserve the character of the random search.

Also, every mass has its speed and acceleration that are shown by $V_i(t)$ and $a_i(t)$ respectively. According to Newton's second law, every mass accelerates at the direction of the exerted force, by the amount directly proportional to the magnitude of the force and inversely proportional to the value of the mass, as expressed by Eq. (26). On the other hand, the velocity of each factor, at the next time, equals the sum of coefficients of the current velocity and the acceleration factor, as given in Eq. (27).

$$a_i^d(t) = F_i^d(t) / M_i(t) \tag{26}$$

$$v_i^d(t+1) = rand \times v_i^d(t) + a_i^d(t) \tag{27}$$

When the acceleration and velocity of each mass are calculated, the new situation of the i^{th} factor, at dimension d, is computed based on Eq. (28).

$$x_i^d(t+1) = x_i^d(t) + v_i^d(t+1) \tag{28}$$

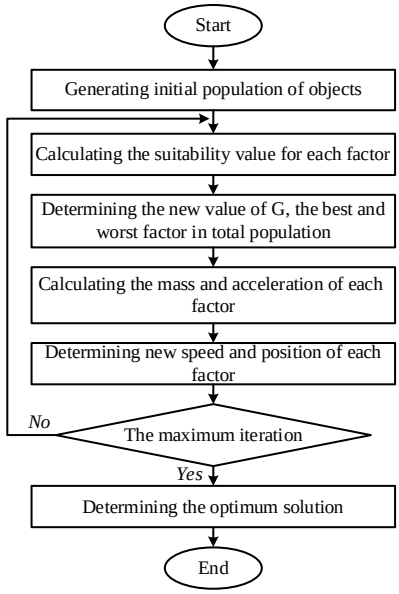


Figure 2. Flowchart of the gravitational search algorithm

The new situations are considered as the places of masses in the searching space and their weights are normalized by equations (29) and (30):

$$m_i(t) = \frac{fit_i(t) - worst(t)}{best(t) - worst(t)} \tag{29}$$

$$M_i(t) = m_i(t) \Big/ \sum_{j=1}^{N} m_j(t) \tag{30}$$

In the above equations: $fit_i(t)$ is the weight of being suitable for mass factor at time t, $worst(t)$ is the weight of the worst population's factor at time t, $best(t)$ is the weight of the best population's factor at time t. The amounts of $worst(t)$ and $best(t)$ in the problem of searching for the minimum point can be calculated based on the equations (31) and (32):

$$worst(t) = \max\{fit_i(t)\} \quad \forall\, i \in \{1,2,....,N\} \tag{31}$$

$$best(t) = \min\{fit_i(t)\} \quad \forall\, i \in \{1,2,....,N\} \tag{32}$$

Stages of Gravitational Search Algorithm (*GSA*) are shown in Fig. 2 [27-28].

IV. Simulation and Result

A. Case study

The system discussed in this paper includes an 18-bus IEEE test distribution system, which is part of the 30-bus IEEE test system with a voltage of 33 kV, which is shown in Fig. 3. This system has four distributed generators units (DGs) in buses 8, 7, 2, and 14 and two uncertain stochastic generation units (USG) in buses 15 and 18. It is expected that the VPP scheme can control the total amount of distributed resources using a VPP strategy to achieve our main function, which is the maximum profit in this article. After the EMS negotiation and agreement with the owners of DGs, the price is determined one day in advance and the technical and economic specifications related to DGs are specified, as the facets for dispatchable units and stochastic generation units are shown in Tables (1) and (2), respectively. EMS for system control and management, has a telecommunication connection with the whole system, to receive data from DG owners or send signals to control the disconnection and connection of DGs. There are also costs associated with control, communication, and coordination that need to be considered [29-30].

Another strategy that COPs use to maximize profits is arbitration so that VPP can buy or sell energy using several points that are connected to the national network. In this system, buying and selling from the market is done through three substations that have different locational marginal prices (LMP), which are shown in buses 1, 11, and 16 in Fig. 3. Because the national network market is based on LMP, and this effect is due to congestion limits, which cause different prices in the three buses. VPP operators take advantage of this situation and buy power from cheaper buses and sell them in more expensive buses. As a result, they can make a significant profit from this position. The projected market price for 24 hours is shown in Fig. 4.

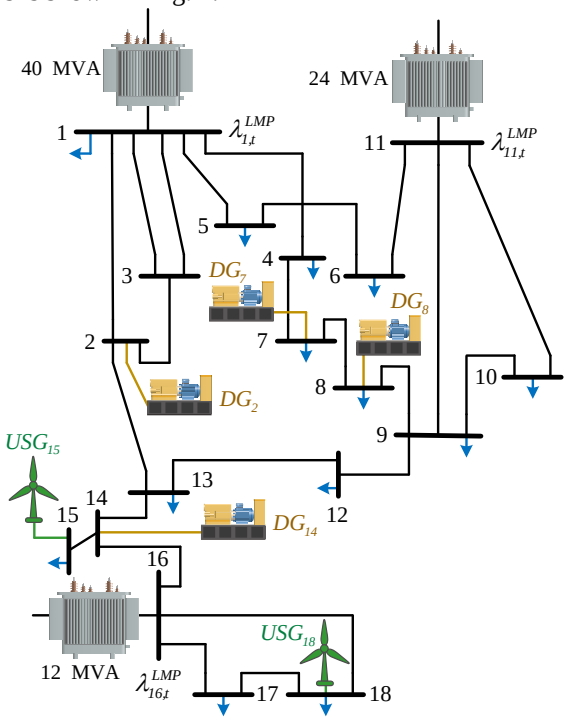


Figure 3. Schematic diagram of 18-bus IEEE test distribution system

TABLE I. SPECIFICATIONS OF DISPATCHABLE UNITS

	DG_2	DG_7	DG_8	DG_{14}
$P_i^{DG.Min}$	0	0	0	0
$P_i^{DG.Max}$	5	6.5	7	9
$\lambda_i^{DG.Cost}$	47	50	44	57
rur_i^{up}	1	1.3	1.45	1.875
rdr_i^{down}	1	1.3	1.45	1.875
$\lambda_i^{DG.Startcost}$	25	25	63	63
$\lambda_i^{DG.Shutcost}$	31	31	31	31
UT_i^{DG}	2	2	2	2
DT_i^{DG}	2	2	2	2

TABLE II. SPECIFICATIONS FOR STOCHASTIC UNITS

	USG_{15}	USG_{18}
$P_i^{USG.Min}$	0	0
$P_i^{USG.Max}$	11.5	9
$\lambda_i^{USG.Cost}$	70	83

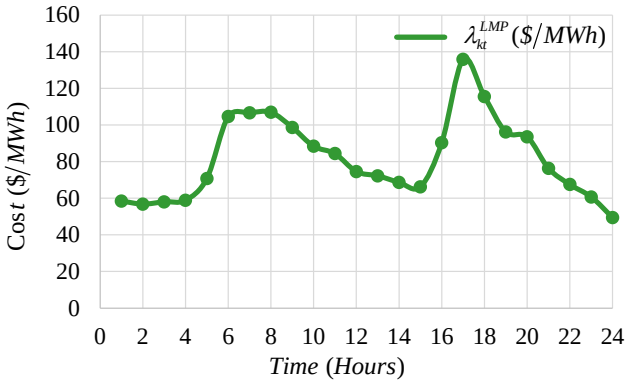


Figure 4. Predicted market energy prices

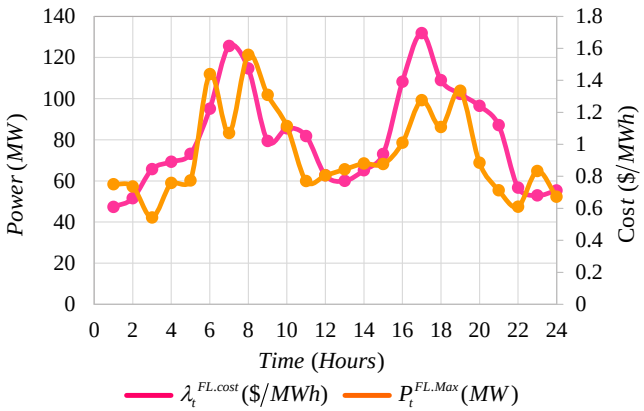


Figure 5. Amount of power and cost associated with flexible loads

In this case study, VPP continues to operate as CVPP. Based on this, the market price in busses 11, 1, and 16 can be predicted at 105, 95, and 100% of the market price of the next day, respectively. The capacity of each transformer in each substation is also shown in Fig. 3.

One of the most profitable features available from a VPP perspective is flexible loads that can give the VPP more flexibility, balancing many RES fluctuations and load shedding during hours when electricity is expensive. The maximum amount of load that can be shed per hour as well as the cost of load shedding is shown in Fig. 5. Another thing that a VPP has to do is to supply the loads inside the VPP, so that the VPP in this case performs its duties as a retailer, as a result, being a retailer with good planning can make a VPP profitable. Therefore, the amount of power required to feed the load of normal subscribers within the VPP and the amount of power required by the bilateral contract must also be specified. To ensure commercial profitability, the CVPP needs to negotiate some bilateral agreements before entering the market. Fig. 8 shows the amount of power contracted for 24 hours, such that the contract price is set at $ 89 per MWh. It should be noted that all decisions are based on the VPP view and DGs owners do not interfere in the decisions and all the uncertainties in RES are due to wind, solar and other environmental factors are the responsibility of DG owners.

B. Discussion and analysis of the results

The calculated profit employing GSA Algorithm in MATLAB software is given by 23528$. Using the analysis of the three grid supply points to the upstream network, with respect to Fig. 6, and the power consumption of normal and contracted subscribers shown in Figs. 7 and 8, it is clear that the amount of energy purchased from buses 1 and 16 is more than the energy consumption of the subscribers in the VPP. In fact, this extra amount of power is sold in bus 11 where the energy price is higher so that the negative values are related to the purchase of energy and the positive values indicate the sale of energy to the energy market.

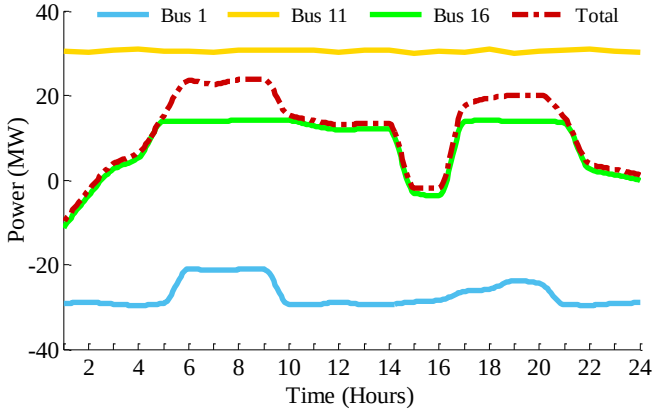


Figure 6. Strategy for buying and selling energy from the national network

Fig. 7 also demonstrates the power consumption of normal subscribers as given in Eq. (21).

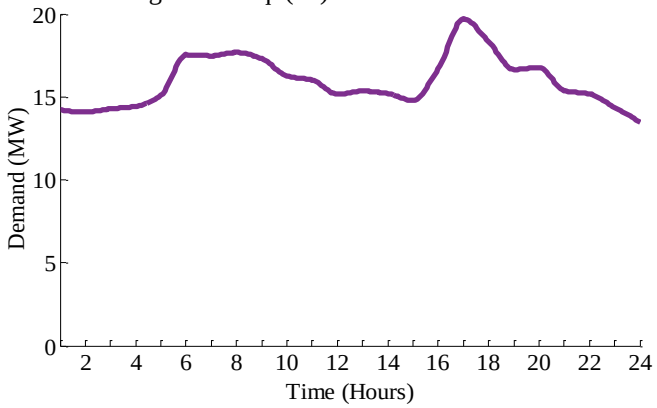


Figure 7. Hourly demand of normal loads

According to Eq. (20), the virtual power plant has to supply power to the contract subscribers in 24 hours, but Eqs. (18) and (19), which represent a 5% deviation in the delivered capacity, give the VPP more flexibility, so that during hours when the market price is higher delivery capacity decreases, whereas

during hours when the market price is lower, it delivers more energy to the contracted subscribers, which can be clearly seen in Fig. 8. Therefore, VPP not only is able to predict the market price precisely when negotiating for a bilateral contract, it can also make more profit through various methods such as the mentioned constraints and also by enabling demand response program developed in [31].

The amount of load shedding from the flexible loads curtailed by the VPP in 24 hours is shown in Fig. 9. Depending on the load shape, the load increases during the hours when the energy price is higher, and vice versa, during the hours when the energy price decreases, demand level will be reduced.

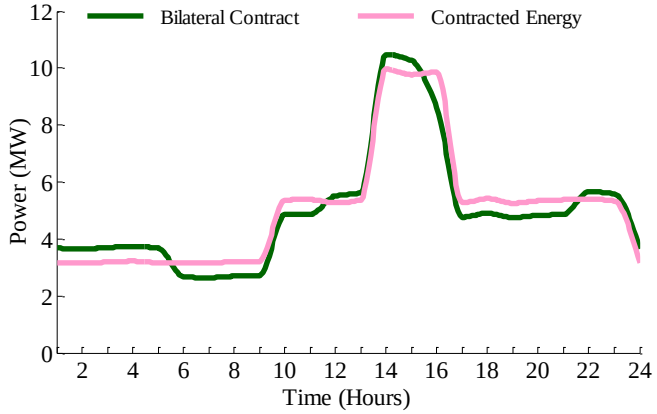


Figure 8. Delivered and contracted power from bilateral contracts

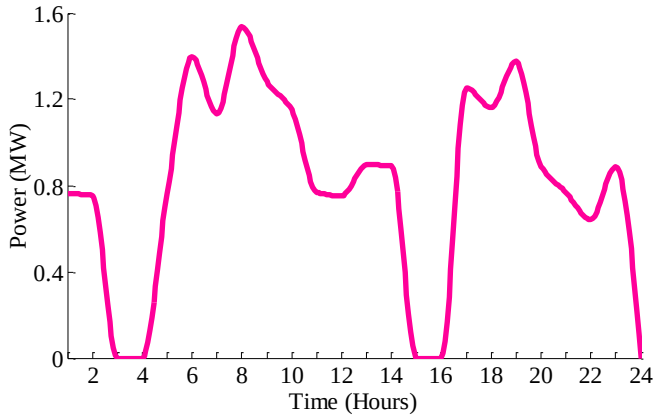


Figure 9. Load shedding amount imposed on flexible loads

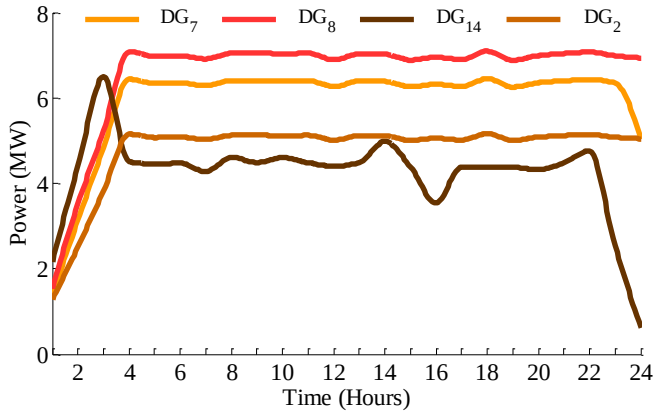


Figure 10. The amount of purchased power from dispatchable DG units

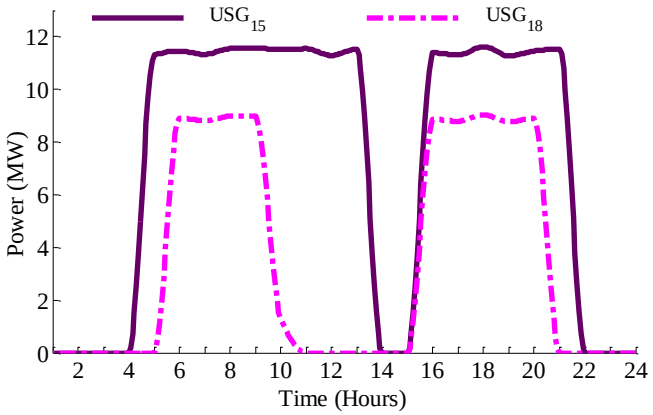


Figure 11. The amount of purchased power from stochastic generation plants with uncertainty

The amount of power purchased from dispatchable units where the commitment schedule is arranged by control signals from VPP operators to the DGs, which is shown in Fig. 10. In addition, Fig. 11 also shows the amount of power purchased from stochastic generation units for 24 hours.

V. Conclusions

Increasing demand for electricity, environmental concerns, and uneconomical development of centralized power plants has inclined the restructured electricity industry planners to integrate distributed generation resources in distribution networks. One of the new concepts that have been proposed to solve the challenges and the possibility of efficiency improvement of distributed generation resources, is creating a structure for the coordinated management and operation of distributed resources under the name of a virtual power plant. In this article, the concept of the virtual power plant scheme as one of the methods of distributed generation integration and providing an optimal model for short-term operation planning of internal resources of the virtual power plant to maximize profits in the 24-hour planning period is discussed. As the results show, when a virtual power plant enters the grid, it undertakes the provision of part of demand satisfaction and increases the total profit. Also, by importing a virtual power plant, in addition to the distributed generation, which by itself increases the reliability of the system, the total profit will increase. On the other hand, as the average load gets closer to the maximum load, the margins of profitability will be expanded.

References

[1] Zhong, W., Chen, J., Liu, M., Murad, M. A. A., & Milano, F. (2021). Coordinated Control of Virtual Power Plants to Improve Power System Short-Term Dynamics. *Energies*, *14*(4), 1182.

[2] Nematollahi, A. F., Shahinzadeh, H., Nafisi, H., Vahidi, B., Amirat, Y., & Benbouzid, M. (2021). Sizing and Sitting of DERs in Active Distribution Networks Incorporating Load Prevailing Uncertainties Using Probabilistic Approaches. *Applied Sciences*, *11*(9), 4156.

[3] Shaneh, M., Shahinzadeh, H., Iranpour, M., Bayindir, R., Nafisi, H., & Marzband, M. (2020, December). Risk Curtailment Assessment in Smart Deregulated Grid with the Presence of Renewable and Storage Sources. In *2020 6th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS)* (pp. 1-8). IEEE.

[4] Sheidaei, F., & Ahmarinejad, A. (2020). Multi-stage stochastic framework for energy management of virtual power plants considering

electric vehicles and demand response programs. *International Journal of Electrical Power & Energy Systems*, *120*, 106047.

[5] Mirhedayati, A. S., Shahinzadeh, H., Nafisi, H., Gharehpetian, G. B., Benbouzid, M., & Shaneh, M. (2021, August). CHPs and EHPs Effectiveness Evaluation in a Residential Multi-Carrier Energy Hub. In *2021 25th Electrical Power Distribution Conference (EPDC)* (pp. 42-47). IEEE.

[6] Maanavi, M., Najafi, A., Godina, R., Mahmoudian, M., & MG Rodrigues, E. (2019). Energy management of virtual power plant considering distributed generation sizing and pricing. *applied sciences*, *9*(14), 2817.

[7] Mirhedayati, A. S., Hadadi, Z., Zanjani, S. M., Shahinzadeh, H., Bayindir, R., Shaneh, M., & Moradi, J. (2021). Optimal operation of combined heat and power in competitive electricity markets: a case study in IAUN. *International Journal of Renewable Energy Research (IJRER)*, *11*(3), 1013-1022.

[8] Wang, X., Liu, Z., Zhang, H., Zhao, Y., Shi, J., & Ding, H. (2019, May). A review on virtual power plant concept, application and challenges. In *2019 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia)* (pp. 4328-4333). IEEE.

[9] Rouzbahani, H. M., Karimipour, H., & Lei, L. (2021). A review on virtual power plant for energy management. *Sustainable Energy Technologies and Assessments*, *47*, 101370.

[10] Sadeghi, S., Jahangir, H., Vatandoust, B., Golkar, M. A., Ahmadian, A., & Elkamel, A. (2021). Optimal bidding strategy of a virtual power plant in day-ahead energy and frequency regulation markets: A deep learning-based approach. *International Journal of Electrical Power & Energy Systems*, *127*, 106646.

[11] Rashidizadeh-Kermani, H., Vahedipour-Dahraie, M., Shafie-Khah, M., & Siano, P. (2021). A stochastic short-term scheduling of virtual power plants with electric vehicles under competitive markets. *International Journal of Electrical Power & Energy Systems*, *124*, 106343.

[12] Rahimi, M., Ardakani, F. J., & Ardakani, A. J. (2021). Optimal stochastic scheduling of electrical and thermal renewable and non-renewable resources in virtual power plant. *International Journal of Electrical Power & Energy Systems*, *127*, 106658.

[13] Behi, B., Arefi, A., Jennings, P., Gorjy, A., & Pivrikas, A. (2021). Advanced monitoring and control system for virtual power plants for enabling customer engagement and market participation. *Energies*, *14*(4), 1113.

[14] Guo, W., Liu, P., & Shu, X. (2021). Optimal dispatching of electric-thermal interconnected virtual power plant considering market trading mechanism. *Journal of Cleaner Production*, *279*, 123446.

[15] Dashtdar, M., Najafi, M., & Esmaeilbeig, M. (2020). Probabilistic planning for participation of virtual power plants in the presence of the thermal power plants in energy and reserve markets. *Sādhanā*, *45*(1), 1-9.

[16] Vahedipour-Dahraie, M., Rashidizadeh-Kermani, H., Shafie-Khah, M., & Catalão, J. P. (2020). Risk-Averse Optimal Energy and Reserve Scheduling for Virtual Power Plants Incorporating Demand Response Programs. *IEEE Transactions on Smart Grid*, *12*(2), 1405-1415.

[17] Naina, P. M., Rajamani, H. S., & Swarup, K. S. (2017, December). Modeling and simulation of virtual power plant in energy management system applications. In *2017 7th International Conference on Power Systems (ICPS)* (pp. 392-397). IEEE.

[18] GM Abdolrasol, M., Hannan, M. A., Hussain, S. M., Ustun, T. S., Sarker, M. R., & Ker, P. J. (2021). Energy Management Scheduling for Microgrids in the Virtual Power Plant System Using Artificial Neural Networks. *Energies*, *14*(20), 6507.

[19] Lin, J., Zhang, S., Yang, B., Li, W., & Yi, Y. (2021, August). Customer-side Energy Management Considering the Availability of Renewable Virtual Power Plants. In *The Sixth International Conference on Information Management and Technology* (pp. 1-5).

[20] Luo, J., Gao, Y., Yang, W., Yang, Y., Zhao, Z., & Tian, S. (2018). Optimal Operation Modes of Virtual Power Plants Based on Typical Scenarios Considering Output Evaluation Criteria. *Energies*, *11*(10), 2634.

[21] Shahinzadeh, H., Nikolovski, S., Moradi, J., & Bayindir, R. (2021, June). A Resilience-Oriented Decision-Making Model for the Operation of Smart Microgrids Subject to Techno-Economic and Security Objectives. In *2021 9th International Conference on Smart Grid (icSmartGrid)* (pp. 226-230). IEEE.

[22] Bagherzadeh, L., Shahinzadeh, H., & Gharehpetian, G. B. (2019, December). Scheduling of distributed energy resources in active distribution networks considering combination of techno-economic and environmental objectives. In *2019 International Power System Conference (PSC)* (pp. 687-695). IEEE.

[23] Bagherzadeh, L., Shahinzadeh, H., Shayeghi, H., & Gharehpetian, G. B. (2019). A short-term energy management of microgrids considering renewable energy resources, micro-compressed air energy storage and DRPs. *International Journal of Renewable Energy Research*, *9*(4).

[24] Shahinzadeh, H., Moradi, J., Gharehpetian, G. B., Fathi, S. H., & Abedi, M. (2018, November). Optimal Energy Scheduling for a Microgrid Encompassing DRRs and Energy Hub Paradigm Subject to Alleviate Emission and Operational Costs. In *2018 Smart Grid Conference (SGC)* (pp. 1-10). IEEE.

[25] Moazzami, M., Ghanbari, M., Moradi, J., Shahinzadeh, H., & Gharehpetian, G. B. (2018). Probabilistic SCUC considering implication of compressed air energy storage on redressing intermittent load and stochastic wind generation. *International Journal of Renewable Energy Research (IJRER)*, *8*(2), 767-783.

[26] Rashedi, E., Nezamabadi-Pour, H., & Saryazdi, S. (2009). GSA: a gravitational search algorithm. *Information sciences*, *179*(13), 2232-2248.

[27] Sabri, N. M., Puteh, M., & Mahmood, M. R. (2013). A review of gravitational search algorithm. *Int. J. Advance. Soft Comput. Appl*, *5*(3), 1-39.

[28] Rashedi, E., Rashedi, E., & Nezamabadi-pour, H. (2018). A comprehensive survey on gravitational search algorithm. *Swarm and evolutionary computation*, *41*, 141-158.

[29] Peik-Herfeh, M., Seifi, H., & Sheikh-El-Eslami, M. K. (2013). Decision making of a virtual power plant under uncertainties for bidding in a day-ahead market using point estimate method. *International Journal of Electrical Power & Energy Systems*, *44*(1), 88-98.

[30] Shafiekhani, M., Badri, A., Shafie-Khah, M., & Catalão, J. P. (2019). Strategic bidding of virtual power plant in energy markets: A bi-level multi-objective approach. *International Journal of Electrical Power & Energy Systems*, *113*, 208-219.

[31] Hooshmand, R. A., Nosratabadi, S. M., & Gholipour, E. (2018). Event-based scheduling of industrial technical virtual power plant considering wind and market prices stochastic behaviors-A case study in Iran. Journal of cleaner production, 172, 1748-1764.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)
March 2-4, 2022, Higher Education Complex of Bam

An Efficient Collaborative Filtering for Recommendation Systems Using Differential Machine Learning

Maryam Nooraei Abadeh[1], Zahra Derakhshandeh[2], Mansooreh Mirzaie[3]

[1] *Department of Computer Engineering, Abadan Branch, Islamic Azad University, Abadan, Iran,* mnooraei@iauabadan.ac.ir

[2] *Computer Science Department, California State University, East BayHayward, USA,*Zahra.derakhshandeh@csueastbay.edu

[3] *Electrical and Computer Engineering Group, Golpayegan college of Engineering, Isfahan University of Technology, Golpayegan, Iran*, m.mirzaei@iut.ac.ir

***Abstract*—Utilizing deep learning, in contrast to standard recommendation models, is capable of successfully capturing non-linear and non-trivial user-item connections and codifying extremely complex abstractions as data representations in higher layers. Differential machine learning as a type of supervised learning trains models based on the twin networks that consider inputs and labels and differentials of labels to inputs. This paper proposes an efficient collaborative filtering method for recommendation systems using differential machine learning, called Differential Collaborative Filtering (DiffCF). DiffCF integrates the cost of derivatives and errors in values, which increases accuracy remarkably. The experiments are performed to train recommender system models on various datasets with different feature dimensions, while a significant reduction in errors and regularization penalties is obtained in comparison to standard machine learning.**

***Index Terms* — collaborative filtering; recommendation system; differential machine learning; deep learning**

I. INTRODUCTION

Nowadays, having access to an extensive amount of data generated by a considerable number of Internet of Things (IoT) devices, autonomous vehicles, and smartphones has led many applications and services to use machine learning techniques for different purposes, including speech recognition, object detection, recommendation, or prediction to improve user experience on the devices.

As an important category of machine learning algorithms, Recommender Systems [1, 2] offer relevant suggestions to the users of online systems. The necessity of designing effective recommender systems has significantly increased due to the plentiful and rapid growth of using the Internet and the competition among service providers in electronic commerce to provide more accurate and relevant information access, predictions, and services. Collaborative Filtering, as one of the fundamental personalized recommender systems, performs automatic filtering/predictions about the interests of the IoT users by aggregating item ratings according to the similarities of the user with other users in a collaborative manner [3].

Generally speaking, machine learning and deep learning algorithms rely on a vast amount of user data to perform the training tasks. This amount of data might not be available on client devices, like user smartphones. As small datasets like user local and personal data can lead to unreliable learning outcomes, it is essential to develop algorithms to deal with such constraints.

Differential Machine Learning (DML) has been recently proposed as an extension of supervised machine learning [4]. The authors in [4] show that training machine learning models can be significantly improved when the gradients of training labels with respect to training inputs are available. Differential Deep differential learning teaches examples of the shape of the target function, resulting in vastly improved accuracy, illustrated with several numerical examples, both idealized and real world. One of the main benefits of DML, in addition to higher accuracy, is its ability to learn effectively from small datasets and scale up to huge ones.

This paper proposes a collaborative filtering recommender method called DiffCF that utilizes the differential machine learning algorithm capable of learning accurate function approximations with remarkable speed and accuracy from small datasets to larger ones in a stable manner. In section 2, we review some basic concepts regarding collaborative filtering recommendation systems as well as differential machine learning. We then continue our discussions by presenting a literature review on recommender systems in section 3. Our

proposed approach will be presented in section 4. Section 5 presents our experiments comparing the standard CF approach and DiffCF. Our experimental results show a notable improvement in accuracy compared to the traditional CF methods for recommender systems. We conclude the work in the last section, where we discuss our future work.

II. BACKGROUND

A. Collaborative Filtering-based Recommendation System

Recommender systems, as an important category of ML algorithms, offer relevant suggestions to the users of online systems. As stated in [2], "It is often necessary to make choices without sufficient personal experience of the alternatives. In everyday life, we rely on recommendations from other people either by word of mouth, recommendation letters, movie and book reviews printed in newspapers ...". Content-based filtering recommender systems perform prediction based on the history (previous rating) of a user. These recommenders require an adequate amount of information regarding items' features, referred to as a well-defined feature matrix, for the prediction about the interest of the user. Features describe the user's tastes and preferences. Collaborative filtering, instead, is a recommender technique that doesn't require items' or users' features to be known in advance. As the result of collecting and analyzing a large amount of information regarding users' preferences, behaviors, and interactions, collaborative filtering predicts what item a user might like according to the user's similarities to the other users. As also stated in [1], in collaborative filtering, "each user contributes to the learning of new features and what features to choose". While in content-based filtering, a user's interest is used to offer an item to that particular user, collaborative filtering aims at offering an item to a user depending on items that were earlier rated by distinct (similar) users.

The idea behind collaborative filtering is to recommend an item to a user who might have not even used the item before. The recommendation happens due to the similarity of the user to other users who like that item. . This technique assumes that the users who agreed before will probably agree in the future.

Considering user preferences, typically, there are explicit and implicit user ratings. In the former, a user directly provides feedback by rating an item. The latter, however, offers users preference indirectly, for example, by considering the number of page views, clicks, checking if an item is added to the wish list of a user, viewed by the user, so on. The data sets in the recommender systems typically contain a set of items and a set of users who have reacted to (some of) the items. This data can be seen in the form of a user-item matrix $R^{N\times M}$ where each row possibly contains the ratings given by a user, and each column contains the ratings received for an item, assuming having N users and M items. This matrix is typically sparse as users might only rate some items.

Due to the sparsity and scalability of matrix R, this matrix will be decomposed into low-dimensional matrices with *latent factors*[1] or features, resulting in less sparsity. This process is called *Matrix Factorization*.

Matrix factorization model maps N users and M items to a joint latent factor low-dimensional (K) space, where user-item interactions are represented as the inner product of matrices $P \in R^{N\times K}$ indicating user-factor matrix and $Q \in R^{K\times M}$ is the item-factor matrix, as shown in equation (3).

The goal is to find P and Q such that once we multiply them back together, the output of matrix R' is the closest approximation of R, while R' is not sparse anymore. In other words, the predicted rating for user u (vector p_u) on item I (vector q_i) will be calculated by:

$$r'_{ui} = p_u^T q_i = \sum_{f=1}^{k} p_{uf} q_{fi} \tag{1}$$

As in many machine learning algorithms, we need to calculate the minimum loss function defined as the squared error such as $R \sim R^0$.

$$\min_{p,q} \sum (r_{ui} - r'_{ui})^2 + regulize(p_u + q_i) \tag{2}$$
$$= \sum (r_{ui} - p_u^T q_i)^2 + \lambda(\sum_u \|p_u\|^2 + \sum_i \|q_i\|^2)$$

where λ is a regulation parameter.

B. Differential machine learning

As an extension of supervised learning, Differential Machine Learning (DML) [4] is proposed where the models are trained based on inputs and also differentials of labels to inputs. The training phase in DML is done on not only labels and input data but also differentials of labels to the input data. By combining with Automatic Adjoint Differentiation (AAD), DML promises an extremely effective approximation in complex machine learning problems. Unlike usual deep learning, which teaches examples the shape of the target function, instances of differential deep learning learn the shape of the target function, resulting in much-improved performance, as demonstrated by various excremental examples, both theoretical and real-world. DML may be used in other machine learning models, such as Principle Component Analysis (PCA) or conventional regression.

Huge and Savine [4] propose a new structure for ML called twin network, which is built by combining feed-forward and back-propagation in an individual network. A twin network as a fundamental construct for differential training is a neural network of twice the depth that can simultaneously predict values and derivatives for twice computation cost.

In the same way, conventional regularization uses small weights in a differential training mode to impose a penalty on incorrect derivatives. DML effectively mitigates over-fitting without creating a bias, compared to traditional regularization. Therefore, there is no tradeoff bias-variance or need to tweak hyper-parameters by cross-validation. The cost function gradients are calculated by back-propagation through the twin

[1] The row or column associated with a specific user and item is referred to as Latent Factor.

network, including the back-propagation part, which Tensor Flow silently conducts as a part of it.

2.2.1 The twin network

The twin network definition is started by its feed forward equations:

$$z_0 = x$$
$$z_1 = g_{l-1}(z_{l-1})w_l + b_l, l = 1, 2, ..., L \quad (3)$$
$$y = z_L$$

Where the row vector including the n_l pre-activation values is shown by $z_l \in R^{n_l}$ in layer l. Fig. 1 shows a twin network with L = 3 and n = n_0 = 3, n = 5, n_2 = 3, n_3 = 1.

For each layer in twin networks in the range [1, L], an activation function $g_{l-1}: R \to R$ is defined which common choices are Softplus, Relu, and Elu. The notation $g_{l-1}(x)$ determines the element-wise application. We denote $w_l \in R^{n_{l-1} * n_l}$ and $b_l \in R^{n_l}$ the weights and biases of layer l.

Back-propagation is the following step, and it is used to calculate the derivatives of a loss function for optimization of weights and biases. We are concerned with the differentials of the predicted value $y = z_L$ wrt the inputs $z_0 = x$ and other variables are obtained by reversing the differentiation of the first lines of Equation 3.

$$\bar{z}_l = \bar{y} = 1$$
$$\bar{z}_{l-1} = \bar{z}_l w_l^T \cdot g'_{l-1}(z_{l-1}), l = 1, ..., L \quad (4)$$
$$\bar{x} = \bar{z}_0$$

Fig. 1 illustrates a twin network with feed-forward and backward parts with corresponding mathematical relations.

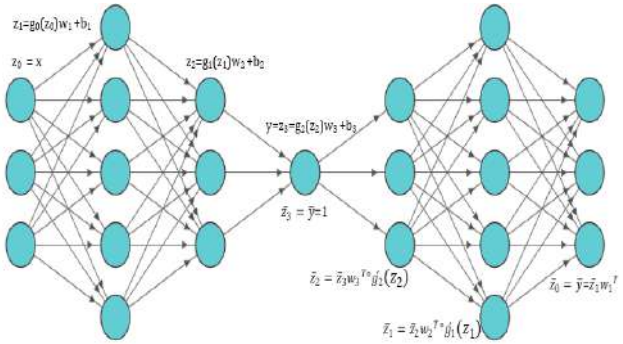


Fig. 1. The twin network [4]

III. RELATED WORK

Recommender systems can estimate users' preference on existing items, which users may like them especially [5, 6]. Recommendation models can be classified into three categories [3]: collaborative filtering and content-based, as well as a hybrid recommender system.

Collaborative filtering recommends by a learning method on user-item previous interactions. These interactions can be in the form of explicit (e.g., the user's previous ratings) or implicit feedback (e.g., browsing history). Content-based recommendation operates based on a comparison based on the auxiliary information of items and users. The supplementary information can include a wide range of data such as texts, images, and videos.

A hybrid model is a kind of recommender system in which two or more types of recommendation strategies can be integrated into that [6,7]. Collaborative filtering (CF) is a popular approach in recommender systems, which can solve many of our problems.

Traditional CF-based approaches use the user-item matrix. This matrix encodes the user's individual preferences for items to learn and complete the recommendation process.

The rating matrix is commonly very sparse in real-world applications, which causes CF-based methods to degrade significantly in terms of recommendation performance. To solve the data sparsity problem, some developed CF methods utilize more additional information.

Since the user-item matrix is sparse and the existence of additional information, the performance would be degrading in learning of the latent factors. Some researchers advanced technics for effective learning, such as a deep-learning collaborative filtering method, a collaborative filtering recommendation algorithm based on a specific model.

Deep learning technology has recently acquired significant progress, and it has made it a hot topic in machine learning and data mining technologies. Deep learning was initially used in image processing and natural language processing fields, producing significant results in these fields [8]. Deep learning constructs a multi-layer, nonlinear, and layer-to-layer interaction network structure, so by using the network structures, this method is approached to a complex multivariate function in terms of a training target. It also learns the basic features from unlabeled training data [9]. Deep learning aims to extract the multi-level features representation from data.

The main idea is extracting features via combining low-level features and, as a result of this, forming the high-level semantic abstractions. Furthermore, deep learning can solve the problem of manually designing features which is a big problem in traditional machine learning. In some cases, containing a large amount of unbalanced data such as image and voice, deep learning would learn all effective features from a large variety of unlabeled training data.

Deep learning has had a significant impact on recommendation models, allowing for more flexibility in recommender implementation. Deep learning-based recommender systems (DLRS) have recently grown strongly to deliver high-quality recommendation systems.

The authors in [10] aim to assess recent deep learning-based recommender system research. They present a taxonomy of deep learning-based recommendation models, as well as a thorough examination of the current state-of-the-art. Finally, they look at current trends and provide new insight into this exciting new development sector.

Traditional recommender methods (matrix factorization, factorization machines, and so on) may be represented as neural/differentiable networks and trained with a framework

like Tensorflow or Pytorch, allowing for faster GPU-accelerated training and free automated differentiation.

Deep learning-based recommendation models are further divided into two groups [10]: (1) neural building blocks for Recommendation. Models are split into eight subcategories in this area, followed by the eight deep learning models stated above: multi-layer perceptron, auto-encoder, convolutional neural network, adversarial networks, recurrent neural network, attentional models, restricted Boltzmann machine, neural autoregressive distribution estimation, and deep reinforcement-learning based recommender system, and (2) deep hybrid recommendation models. Some recommendation models based on deep learning employ multiple deep learning techniques. Twin networks used in DML are based on the universal approximation theorem, so we focus on Multilayer feed-forward networks [11].

Alashkar et al. [12] presented a makeup suggestion algorithm based on MLPs. This study model labeled instances and expert rules using two identical MLPs. By reducing the differences between their outputs, the parameters of these two networks are changed concurrently. Cheng et al. [13] presented a comprehensive and deep model for an App recommender system for recommending users utilizing Google Play. For video recommendation on YouTube, Covington et al. [14] introduced a deep neural network-based recommendation system by applying multilayer perceptron.

The authors of [15] utilized an MLP to estimate the interaction function, and they found that it outperformed standard techniques like MF. They offer when trained with momentum-based gradient descent, and neural models outperform better than traditional machine learning models like BPR, MF, and CML on interaction-only data [16]. These models, however, may also be considered neural architectures because they make use of recent deep learning breakthroughs such as Dropout, Batch Normalization, and Adam [15, 17].

The authors in [18] utilized a factorization machine as a multi-criteria linear family member. They used MLP to replace second-order interactions to propose a recommended model using dropout and batch normalization to regularize the model. By integrating prediction errors in various deep learning layers, the approach proposed in [18] uses the dependability concept's potential to improve prediction and recommendation quality.

Rendle et al. [20] examine the practical problems occurring when using models learned similarities by MLPs. They show that MLPs are too expensive to employ in production contexts for an item recommendation. In contrast, dot products allow for highly efficient retrieval methods to be used.

The authors in [21] proposed to provide a semi-supervised learning approach to relieve data scarcity by smoothing between nearby users and POIs (points of interest) and address diverse contexts by the regularizing preferences of users based on context graphs.

Lian et al. presented [22] a new Compressed Interaction Network (CIN) to explicitly generate feature interactions on a vector-by-vector basis. We show that recurrent neural networks and convolutional neural networks share some functions with CINs.

IV. THE PROPOSED APPROACH

This paper proposes differential learning as an approach that enables differential model training on a broad set of decentralized data. The technique promises vastly improved accuracy in model training by using twin networks.

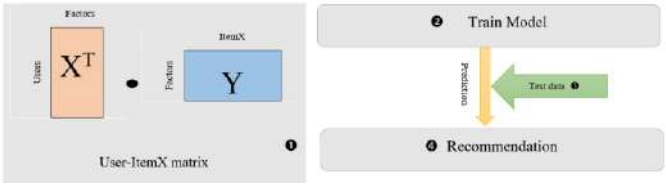


Fig. 2. The proposed architecture for DiffCF

The training is performed by calculating labels that are obtained by a function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ which the network is defined according to weights and biases in the form of $\hat{f}(x; \{w_l, b_l\}_{l=1,..,L})$. The training set consists of *m* sample user vectors $x_1, x_2, x_3, \ldots, x_m$ with labels

$$y_1 = f(x_1), \ldots, y_m = f(x_m). \tag{5}$$

We also provide derivatives labels:

$$\overline{x_1} = \frac{\partial f(x_1)}{\partial x_1}, \ldots, \overline{x_m} = \frac{\partial f(x_m)}{\partial x_m}. \tag{6}$$

Differential training for local user data consists of finding weights and biases such that the network predictions on the function compute instances that are close to the labels f. The weights and biases are optimized by minimizing a cost function:

$$\min_{wl,bl} c(\{w_l, b_l\}_{l=1,\ldots,L})$$

The mean squared error in values is the standard regression cost, with no penalty for incorrect derivatives. The cost of errors in values and derivatives is combined in differential training:

$$f(\{w_l, b_l\}_{l=1,\ldots,L}) = \frac{1}{m} \times (Z_L - Y)^T (Z_L - Y) + \lambda \frac{1}{m} + tr[\beta^T (\overline{Z_0} - \overline{X})^T (\overline{Z_0} - \overline{X})\beta] \tag{7}$$

where β is the diagonal matrix of the weights and tr is the trace of a matrix. The other parameters are defined as follows:

$$x = \begin{bmatrix} x_1 \\ , \\ , \\ , \\ x_m \end{bmatrix} \in R^{m \times n} \ldots\ldots \quad y = \begin{bmatrix} y_1 \\ , \\ , \\ , \\ y_1 \end{bmatrix} \in R^m \quad \overline{x} = \begin{bmatrix} \overline{x1} \\ , \\ , \\ , \\ \overline{xm} \end{bmatrix} \in R^{m \times n}$$

$$z_l = \begin{bmatrix} z_l^{(1)} \\ , \\ , \\ , \\ z_l^{(m)} \end{bmatrix} \in R^{m \times n_l} \qquad \bar{z}_l = \begin{bmatrix} \bar{z}_l^{(1)} \\ , \\ , \\ , \\ \bar{z}_l^{(m)} \end{bmatrix} \in R^{m \times n_l} \tag{8}$$

The following section presents efficient steps for collaborative filtering using differential machine learning. Fig. 2 shows the architecture of the proposed recommendation system using a differential machine learning model.

According to the proposed architecture in Fig.2, we consider the case of explicit feedback collaborative filtering. Our Differential Collaborative Filtering (DFCF) follows:

– The user factor vectors p_u, $1 \le u \le N$ are updated on each client u, using its local data, and q_i, $1 \le i \le M$ is calculated as the average values of local user factors. The updates for each user u can be performed independently of other users and using client u's local data.

– To learn the item factor vectors q_i, the interaction of all users with the item i is needed. Therefore, clients can not directly calculate the update of q_{ui} parameters, and instead, this process will be performed on the server.

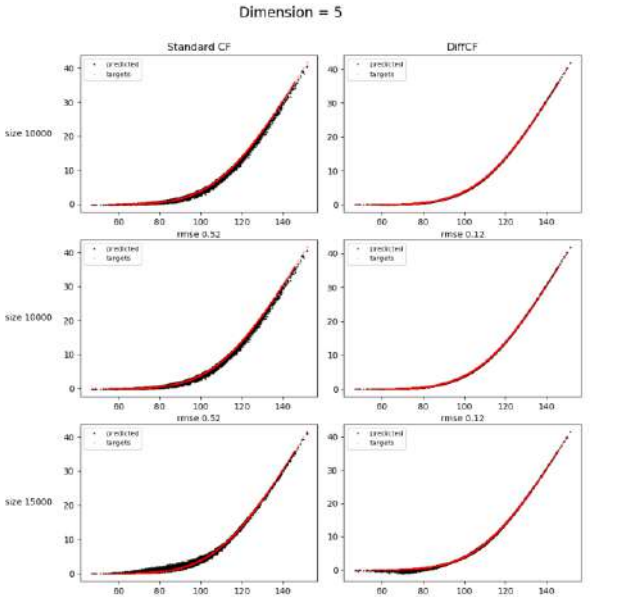
– In addition to what we discussed above, the cost function optimizing all users u and items i reflects the differential learning.

$$J(P,Q) = \frac{1}{m}\sum_u \sum_i ((r_{ui}^L - p_u^T q_i)\beta)^2 + \lambda \frac{1}{m}\left(\sum_u tr[((\overset{-0}{r}_{ui} - \bar{p}_u)\beta]^2 + \sum_i tr[((\overset{-0}{r}_{ui} - \bar{q}_i)\beta]^2\right) \tag{9}$$

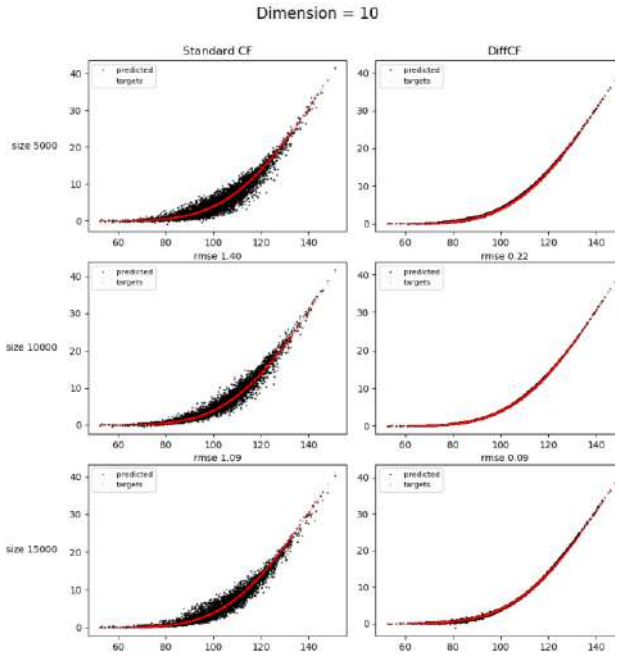
– Finally, differential learning is utilized to build DiffCF as an efficient approach for recommender systems.

V. EXPERIMENTAL RESULTS

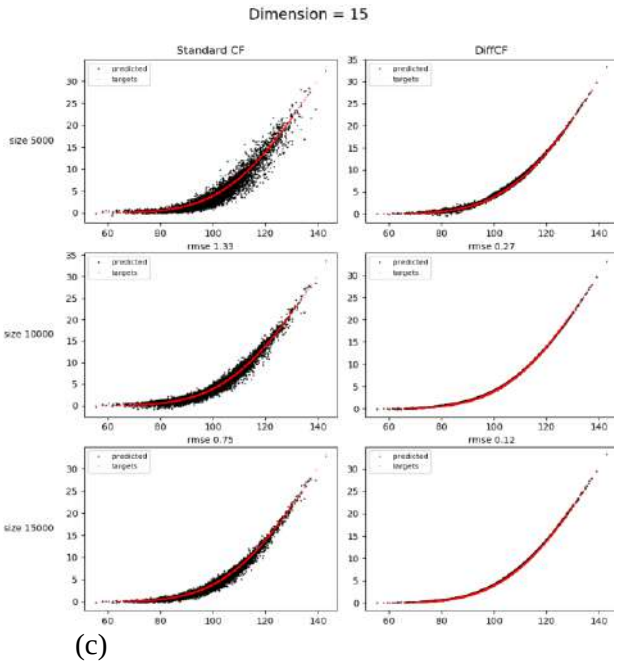
Using the Lonstaff-Schwartz [23] method, we learn from samples. Our inputs are simulated Markov states, labels are payoffs chosen from the same Monte-Carlo route, and differential labels are pathwise derivatives typically generated using AAD.



(a)



(b)



(c)

Fig. 3. The comparison between DiffCF and standard CF

Fig. 3 (a), (b), and (c) show the results of training a number of networks on 5000, 10000, and 15000 samples. For each dataset, three-dimension sizes 5, 10, and 15 are generated to forecast values and derivatives in various scenarios using standard regression models compared to the correct results computed using differential machine learning. As we can see, the twin network outperforms better utilizing differential labels and a substantial improvement for derivative values. It discovered, in particular, that the delta value in differential labels leads to increase training accuracy according to the functions of the forward and backward training.

The classical network doesn't learn this property well, even with small datasets, e.g., 5000 instances. It leads to under-fits in training data and predicts different values or deltas while the twin network overcomes this with AAD pathwise derivatives. Even with 15,000 data, the output training results of the standard ML network are worse than that of the twin network with 5,000 data. As a result, twin networks are suitable for applying to critical real-world concerns like laws and risk reports with many scenarios.

VI. CONCLUSION AND FUTURE DIRECTIONS

Deep learning-based recommender systems have built more suitable models to improve accuracy and user satisfaction in recent years. In this paper, DiffCF has proposed to use twin networks in the differential learning process. As a result, performance and accuracy will significantly improve model training. The empirical evidence of DiffCF implementation on three different training recommendation system networks proves increasing training accuracy utilizing twin networks. As a practical outcome, the results show that DiffCF can provide a more qualified recommender system than standard machine learning methods. We consider DiffCF as the first step toward recommendation systems with a differential model. We will examine the usage of DiffCF via empirical results on both synthetic and real-world recommender system datasets. We also vary the number of hidden layers to investigate deeper neural networks' accuracy and time complexity.

REFERENCES

[1] Y. Hu, Y. Koren, and C. Volinsky, "Collaborative filtering for implicit feedback datasets," in *2008 Eighth IEEE International Conference on Data Mining*, 2008, pp. 263-272: Ieee.
[2] Resnick, Paul, and Hal R. Varian. "Recommender systems." *Communications of the ACM*, vol. 40, no. 3, pp. 56-58, 1997.
[3] D. Jannach, M. Zanker, A. Felfernig, and G. Friedrich, *Recommender systems: an introduction*. Cambridge University Press, 2010.
[4] B. N. Huge and A. Savine, "Differential Machine Learning," *Available at SSRN 3591734*, 2020.
[5] H. Abdollahpouri, E. C. Malthouse, J. A. Konstan, B. Mobasher, and J. Gilbert, "Toward the next generation of news recommender systems," in *Companion Proceedings of the Web Conference 2021*, 2021, pp. 402-406.
[6] F. Ricci, L. Rokach, and B. Shapira, "Recommender systems: introduction and challenges," in *Recommender systems handbook*: Springer, 2015, pp. 1-34.
[7] R. Burke, "Hybrid recommender systems: Survey and experiments," *User modeling and user-adapted interaction*, vol. 12, no. 4, pp. 331-370, 2002.
[8] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," N*ature*, vol. 521, no. 7553, pp. 436-444, 2015.
[9] Y.-x. Peng *et al.*, "Cross-media analysis and reasoning: advances and directions," *Frontiers of Information Technology & Electronic Engineering*, vol. 18, no. 1, pp. 44-57, 2017.
[10] S. Zhang, L. Yao, A. Sun, and Y. Tay, "Deep learning based recommender system: A survey and new perspectives," *ACM Computing Surveys (CSUR)*, vol. 52, no. 1, pp. 1-38, 2019.
[11] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural networks*, vol. 2, no. 5, pp. 359-366, 1989.
[12] T. Alashkar, S. Jiang, S. Wang, and Y. Fu, "Examples-rules guided deep neural network for makeup recommendation," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2017, vol. 31, no. 1.
[13] H.-T. Cheng *et al.*, "Wide & deep learning for recommender systems," in *Proceedings of the 1st workshop on deep learning for recommender systems*, 2016, pp. 7-10.
[14] P. Covington, J. Adams, and E. Sargin, "Deep neural networks for youtube recommendations," in *Proceedings of the 10th ACM conference on recommender systems*, 2016, pp. 191-198.
[15] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T.-S. Chua, "Neural collaborative filtering," in *Proceedings of the 26th international conference on world wide web*, 2017, pp. 173-
[16] Y. Tay, L. Anh Tuan, and S. C. Hui, "Latent relational metric learning via memory-based attention for collaborative ranking," in *Proceedings of the 2018 World Wide Web Conference*, 2018, pp. 729-739.
[17] S. Zhang, L. Yao, A. Sun, S. Wang, G. Long, and M. Dong, "Neurec: On nonlinear transformation for personalized ranking," *arXiv preprint arXiv:1805.03002*, 2018.
[18] X. He and T.-S. Chua, "Neural factorization machines for sparse predictive analytics," in *Proceedings of the 40th International ACM SIGIR conference on Research and Development in Information Retrieval*, 2017, pp. 355-364.
[19] J. Bobadilla, S. Alonso, and A. Hernando, "Deep Learning Architecture for Collaborative Filtering Recommender Systems," *Applied Sciences*, vol. 10, no. 7, p. 2441, 2020.
[20] S. Rendle, W. Krichene, L. Zhang, and J. Anderson, "Neural collaborative filtering vs. matrix factorization revisited," in *Fourteenth ACM Conference on Recommender Systems*, 2020, pp. 240-248.
[21] C. Yang, L. Bai, C. Zhang, Q. Yuan, and J. Han, "Bridging collaborative filtering and semi-supervised learning: a neural approach for poi recommendation," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2017, pp. 1245-1254.
[22] J. Lian, X. Zhou, F. Zhang, Z. Chen, X. Xie, and G. Sun, "xdeepfm: Combining explicit and implicit feature interactions for recommender systems," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2018, pp. 1754-1763.
[23] Longstaff, F. A., & Schwartz, E. S. (2001). Valuing American options by simulation: a simple least-squares approach. The review of financial studies, 14(1), 113-147.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)
March 2-4, 2022, Higher Education Complex of Bam

An efficient outlier detection method based on distance ratio of k-nearest neighbors

Reza Heydari Gharaei[1], Rasoul Sharify[2], Hossein Nezamabadi-pour[3]

[1]*Department of Electrical Engineering, Shahid Bahonar University of Kerman, Kerman, Iran*, reza.heydari@eng.uk.ac.ir
[2]*Faculty of Engineering Science, College of Engineering, University of Tehran, Tehran, Iran*, rasoul.sharifi@ut.ac.ir
[3]*Department of Electrical Engineering, Shahid Bahonar University of Kerman, Kerman, Iran*, nezam@uk.ac.ir

***Abstract*—Outlier detection is an important topic, which has been investigated in different theoretical and applied sciences. So far, numerous methods and algorithms have been proposed for outlier detection; each of these methods has found a use in some fields and has advantages and disadvantages. The present study presents an efficient method for detecting outliers based on the distance ratio of the *k*-nearest neighbors. In this method, both criteria of density and distance between objects have been included. Contrary to a large number of the previous methods, the density of each object is measured in relation to each neighbor independently, and after considering the distance criterion, the outlierness score of each object is measured collectively. Another feature of this method is its resistance to the *k* parameter. The proposed method has been evaluated in two-dimensional synthetic and multi-dimensional real datasets and compared with other significant algorithms in this field. The results of the experiments proved the efficiency of the proposed method.**

***Index Terms* —Outlier detection; distance-based method; density-based method; outlier; k-nearest neighbor**

I. INTRODUCTION

Data mining is the science of extracting knowledge from datasets through smart methods. One of the first steps toward an integrated analysis of data is outlier detection. Although outliers are often considered anomalies and faults, they may carry important information [1]. The issue of detecting outliers in a dataset is about finding objects that are inconsistent with the normal behaviors of the objects [2]. Hawkins defines an outlier as an observation that significantly deviates from other observations and creates suspicions [3]. Outliers are generally divided into two groups of global and local outliers. A global outlier is an object with a relatively large distance from other objects, while a local outlier normally exists around dense clusters and its near neighbors are being a denser region. In other words, for this type of object, only the ratio of density matters not their distance from other objects. In [4], the difference between these two types of objects is properly explained. Outlier detection has many applications in various fields. Fraud detection in the use of credit cards, medical and health care, detection of an intrusion in cyber security, detection of faults in security systems, and military surveillance on enemy's activities are some uses of these algorithms [5]. In recent years, several methods have been introduced for outlier detection. One of the most extensive classifications of outlier detection methods is presented in [6]. These methods are generally divided into two main groups, namely supervised and unsupervised methods. In supervised methods, some information about the dataset is available, while in unsupervised methods, no information about the dataset is available; consequently, unsupervised methods have been being used more often lately [7]. Two major groups of unsupervised methods are distance-based and density-based methods; both these methods often use the k nearest neighbors of the objects. Some distance-based methods consider a constant radius d for each object and using the number of objects around the sample object within this radius, achieve a criterion for outlierness [8]. Other distance-based methods consider a constant number of nearest neighbors (kNN) for each object, and achieve a criterion for outlierness by measuring the distance between each object and its k nearest neighbors. References [9]–[11] mostly use this method for global outlier detection. In any case, distance-based methods merely detect global outliers since as mentioned before global outliers are farther compared

with other objects; on the other hand, density-based methods, such as [4], [12] mostly focus on local outlier detection. Each method has its own advantages and disadvantages; for instance, although RDOS [12] has a favorable precision, it has a relatively high time complexity since it needs to detect direct, reverse, and shared neighborhoods. Moreover, in LOF [4], if two dense and sparse clusters are next to one another, normal objects on the margins of the sparse cluster will be mistaken for outliers; this shortcoming was to some extent overcome in our previous study (OFIN [13]), using reverse neighborhood. Another method analyzed in this study is LoOP [14]. This method, which is a statistical method for local outlier detection, is significantly sensitive to k, and has low robustness; moreover, it is incapable of detecting global outliers. One of the latest outlier detection methods is MOD [11]; in spite of its high speed and efficiency in global outlier detection, it does not perform well in local outlier detection. This study aimed to combine distance-based and density-based methods to detect global and local outliers, and decrease the sensitivity to k without increasing time complexity and improve the shortcomings of the previous methods. Therefore, the proposed method introduces a more efficient algorithm by presenting a new approach. In the second section of the study, the proposed algorithm is explained; in the third section, the different experiments on synthetic and real datasets using the proposed algorithm and other significant algorithms and their results are compared; finally, in the fourth section, a conclusion is drawn.

II. Proposed Method

The proposed method is based on the k Nearest Neighbor; therefore, first, we need to find the k nearest neighbor of each object. To explain the proposed method, let us consider dataset D with n data. To measure the outlierness of object p ($p \epsilon D$), first, we find its k nearest neighbor.

$$N_k(p) = \{o_i | d(p, o_i) \leq d(p, o_k), 1 \leq i \leq k\} \quad (1)$$

In this equation, o_i is the i-th neighbor of the object p and $d(p, o_i)$ is the distance between p and o_i. For the sake of simplicity, we equate $d(p, o_i)$ with $i\_distance(p)$. In the next step, we need to measure $i\_distance(o_i)$ for $1 \leq i \leq k$. An example of a two-dimensional dataset with 7 objects and assuming $k = 3$ is presented in Fig. 1 for further explanation. o_1, o_2 and o_3 are the first, second, and third neighbor of p, respectively; furthermore, in this figure, o_{1_1}, p and o_{3_3} are the first neighbor o_1, the second neighbor o_2, and the third neighbor o_3, respectively. In this figure, $i\_distance(p)$ is represented with straight red lines and $i\_distance(o_i)$ with straight blue lines. Now, we need to measure the density ratio of each neighbor of p to the object p. To this end, we should calculate $\frac{i\_distance(o_i)}{i\_distance(p)}$. The smaller the result of this fraction is the denser the i-th neighbor of p is in relation to p; therefore, p seems locally more outlier. Since this criterion is density-based, it is unable to detect global outliers and mostly detects local outliers. To detect global outliers, another criterion based on distance

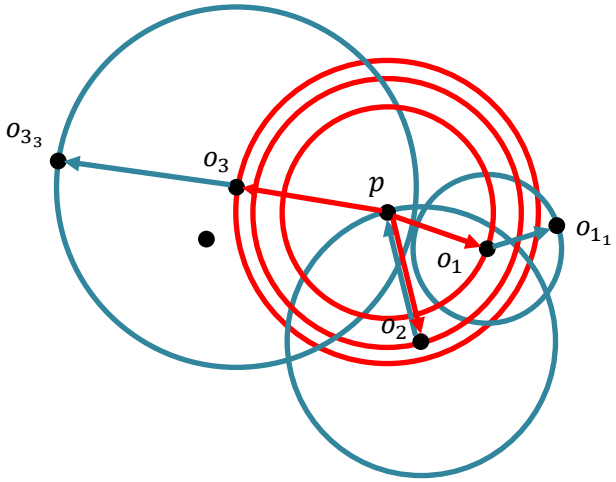


Fig. 1. An example for $i\_distance(p)$ and $i\_distance(o_i)$

must be added to the aforementioned criterion. As we know, the farther p is from its neighbors the higher global outlierness score it has. Therefore, the smaller the result of $\frac{1}{i\_distance(p)}$ is, p is globally more outlier. As seen, this criterion is exclusively based on distance and effectively detects global outliers. Eventually, by combining density and distance and using the resulting aforementioned fractions, we achieve Eq. (2).

$$DR - kNN_i(p) = \frac{i\_distance(o_i)}{i\_distance(p)^2} \quad (2)$$

In fact, Eq. (2), which is the product of the two aforementioned criteria, merely measures the outlierness of p in relation to its i-th neighbor. To measure the outlierness of p, we need to calculate Eq. (2) for all neighbors of p. To this end, based on Eq. (3), the average of the sum of Eq. (2) for all neighbors is calculated.

$$DR - kNN(p) = \frac{1}{k} \times \sum_{i=1}^{k} DR - kNN_i(p) \quad (3)$$

According to Eq. (3), $DR - kNN(p)$ expresses the outlierness of object p. The proposed method that called Distance Ratio of k Nearest Neighbor (DR-kNN), similar to most methods in this field, assigns an outlierness score to each object of dataset D, and at the end, we can discover the most outlier objects by ordering these scores. The smaller the value of $DR - kNN(p)$ is, the higher the outlierness score of p will be, and the greater it is, the more normal p will be. For more simplicity, the aforementioned method is included in Algorithm1 as a pseudocode. Based on this pseudocode, we can consider q objects with smallest values of $DR - kNN$ as outliers. One of the major challenges of kNN-based algorithms is selecting a proper k; therefore, a method with the lowest sensitivity to this parameter is more efficient. One of the advantages of the the proposed method is that it has overcome this challenge to a great extent. The final parts of the next section will cover this topic.

Algorithm 1. $DR-kNN(p)$

Input: A given data set D, number of nearest neighbors k, and number of outliers q

Output: q objects with the lowest $DR-kNN(p)$

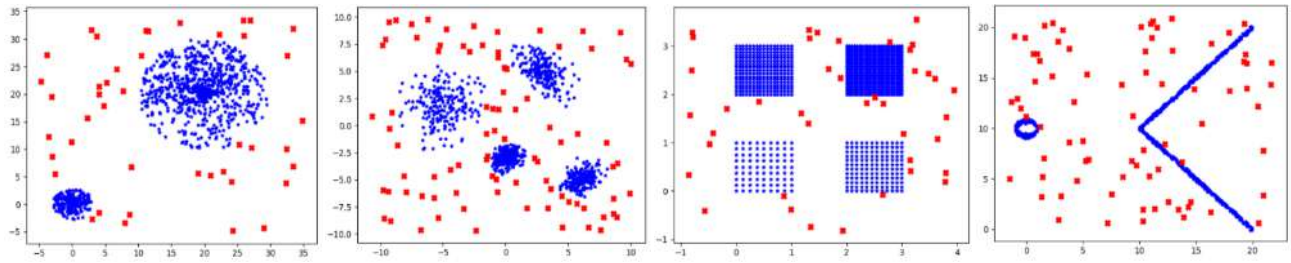
(1) for each object $p \epsilon D$ do
(2) $N_k(p) \leftarrow$ find k-nearest neighbors p
(3) $i \leftarrow 1$
(4) $DR-kNN(p)=0$
(5) for each object $o_i \epsilon N_k(p)$ do
(6) calculate $i\_distance(p)$
(7) $o_{i_i} \leftarrow$ find i-th nearest neighbor of o_i
(8) calculate $i\_distance(o_i)$
(9) $DR-kNN_i(\mathrm{p}) = \frac{i\_distance(o_i)}{i\_distance(p)^2}$
(10) $DR-kNN(p) \mathrel{+}= DR-kNN_i(\mathrm{p})$
(11) end for
(12) $DR-kNN(p) = \frac{1}{k} \times DR-kNN(p)$
(13) end for
(14) select q objects with the lowest $DR-kNN(p)$ to be considered as outliers

III. Experimental Results

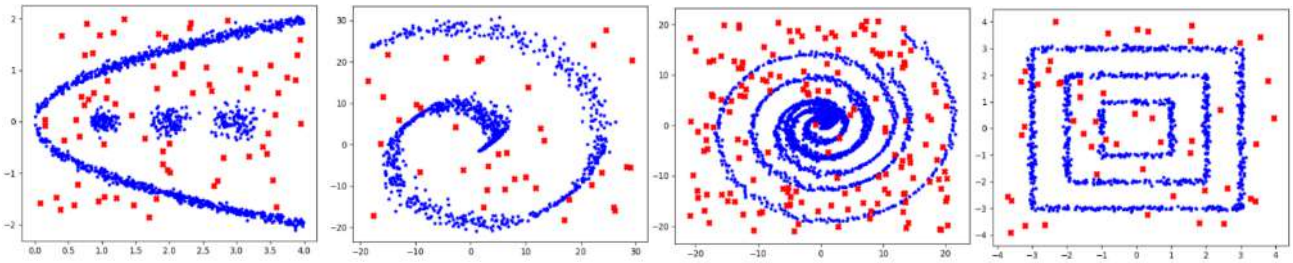
The algorithm was experimented on 8 two-dimensional synthetics and 8 real datasets. Information on all datasets is presented in Table Ⅰ, and the two-dimensional synthetic datasets are given in Fig. 2. Two popular criteria of Precision and AUC [15] were evaluated in the experiments. Using the two aforementioned criteria, the proposed algorithm was compared with LOF, LoOP, RDOS, OFIN, and MOD, which are some of the most popular methods in this field. The precision graphs of all datasets for $5 \leq k \leq 100$ are depicted in Fig. 3. As seen in these graphs, the proposed method has significant advantages in (c), (e), (h), (j), (m), (n), and (p) datasets; moreover, it has a competitive advantage in (a), (d), (f), (k), and (l) datasets, but for large k in (b) and (g) datasets, it is less efficient compared with LOF and RDOS, respectively, and in (i) dataset, it is less efficient compared whit OFIN and LOF. In dataset (o), our method ranked second after MOD algorithm. The mean of AUC in the same range for all datasets is presented in Table Ⅱ. As seen, the proposed method ranked first in all datasets except for dataset3, dataset6, and cardio. Furthermore, for this criterion (AUC), the standard deviation for $5 \leq k \leq 100$ was calculated for all datasets and algorithms. As we know, the smaller the value of standard deviation is, the less the changes of AUC in different k are. In other words, a decrease in this criterion

TABLE Ⅰ. DESCRIPTIONS OF THE 16 USED DATASETS IN EXPERIMENTS

Dataset name	Dimensions	Number of samples	Number of outliers
Dataset 1	2	1043	43
Dataset 2	2	1000	85
Dataset 3	2	1039	41
Dataset 4	2	876	77
Dataset 5	2	1372	72
Dataset 6	2	1037	37
Dataset 7	2	2259	159
Dataset 8	2	1242	50
Annthyroid	6	7200	534
Breastw	30	569	212
Cardio	21	1831	176
Diabetes	8	768	268
Lympho	18	148	6
Optdigits	64	5216	150
Pendigits	16	6870	156
Satellite	36	6435	2056



(a) Dataset 1 (b) Dataset 2 (c) Dataset 3 (d) Dataset 4



(e) Dataset 5 (f) Dataset 6 (g) Dataset 7 (h) Dataset 8

Fig. 2. Two-dimensional synthetic datasets (normal: ★, outlier: ✖)

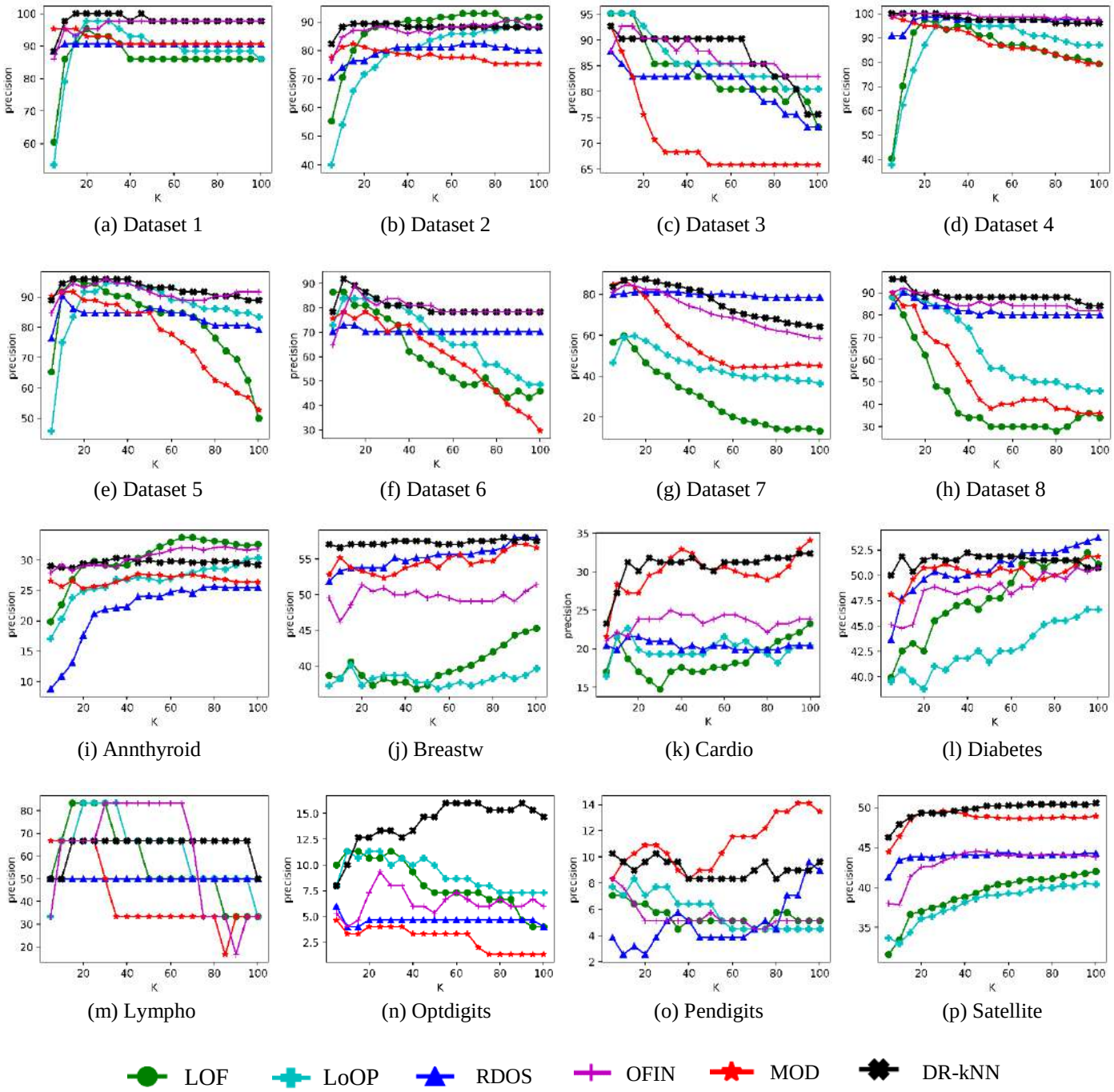


Fig. 3. Detection performance (Precision) of 5 methods on 16 datasets.

indicates a decrease in the sensitivity of the algorithm to k, which in turn increases the robustness of the algorithm. In Table Ⅲ, this criterion is investigated for AUC. As seen, in most datasets, the proposed method has significantly smaller values of standard deviation compared with the other methods, showing its high robustness and efficiency. To analyze the reason for the low standard deviation of the proposed method, we need to heed Eq. (2). As seen in the denominator of the fraction, there is a square of an expression of distance, while there is a first degree of a similar expression in the numerator; therefore, with an increase in k, the value of this fraction decreases in each step. As a result, farther neighbors have a less significant impact on outlierness measurement compared with nearer neighbors, and they lead to a higher stability of the algorithm. According to the results of the experiments, the proposed method is efficient, and can be employed for different types of data.

TABLE Ⅱ. AVERAGE OF AUC IN RANGE $5 \leq k \leq 100$

AUC	LOF	LoOP	RDOS	OFIN	MOD	DR-kNN
Dataset 1	99.06	99.14	98.3	99.97	99.82	**99.99**
Dataset 2	98.34	97.07	97.79	99.54	98.67	**99.63**
Dataset 3	99.01	**99.49**	99.11	99.17	97.64	99.3
Dataset 4	95.32	96.78	98.97	99.91	99.21	**99.98**
Dataset 5	98.12	98.76	99	99.86	98.14	**99.91**
Dataset 6	89.91	96.69	97.83	97.89	95.24	**98.88**
Dataset 7	80.26	79.9	95.31	97.3	89.47	**98.04**
Dataset 8	76.38	82.75	96.72	98.72	77.72	**99.54**
Annthyroid	72.01	74.93	61.48	75.08	70.95	**75.62**
Breastw	57.42	52.58	68.8	70.07	70.92	**75.43**
Cardio	68.49	61.02	55.12	71.53	**76.64**	71.25
Diabetes	64.05	58.63	66.88	67.01	67.3	**69.45**
Lympho	96.36	97.82	96.36	97.24	95.38	**98.33**
Optdigits	50.21	56.21	**60.92**	54.14	42.88	47.32
Pendigits	46.96	48.34	63.19	68.31	67.65	**71.61**
Satellite	55.68	54.75	59.59	62.28	66.27	**68.02**
Average	77.97	78.43	82.21	84.88	82.12	**85.77**

TABLE Ⅲ. STANDARD DEVIATION OF AUC IN RANGE $5 \leq k \leq 100$

σ (AUC)	LOF	LoOP	RDOS	OFIN	MOD	DR-kNN
Dataset 1	1.32	2.93	1.58	0.07	0.13	**0.03**
Dataset 2	4.29	5.97	3.67	0.28	0.37	**0.12**
Dataset 3	0.86	0.34	0.4	0.45	1.19	**0.3**
Dataset 4	5.36	7.29	2.61	0.1	0.65	**0.02**
Dataset 5	1.78	2.82	1.06	0.09	1.3	**0.06**
Dataset 6	6.53	2.4	0.74	1.11	2.81	**0.52**
Dataset 7	2.33	6.27	1.08	1.2	4.93	**0.91**
Dataset 8	11	14.3	1.14	0.49	16	**0.3**
Annthyroid	**1.31**	1.62	4.88	1.94	2.24	1.71
Breastw	4.96	0.64	3.44	1.64	**0.6**	0.72
Cardio	12	4.47	**0.49**	6.09	6.3	5.56
Diabetes	3.72	1.71	3.77	1.67	1.11	**0.53**
Lympho	3.01	1.54	0.87	2.03	1.97	**0.66**
Optdigits	8.07	3.82	2.04	**0.58**	3.69	0.92
Pendigits	**2.56**	2.77	13.1	3.88	4.11	3.03
Satellite	2.29	2.21	**0.7**	2.2	1.43	1.59
Average	4.46	3.82	2.6	1.49	3.05	**1.06**

IV. CONCLUSION

Outlier detection has several applications in various fields. In this study, we aimed to achieve an efficient method for detecting anomalies. The proposed algorithm includes both density and distance is measuring outlierness; therefore, it can detect both global and local outliers. Moreover, the proposed method is resistant to changes in k; consequently, outlierness measurement is performed with a higher precision, which is confirmed by different experiments.

REFERENCES

[1] J. H. VICTORIA and A. JIM, "A Survey of Outlier Detection Methodologies," *Artif. Intell. Rev.*, vol. 22, pp. 85–126, 2004

[2] F. E. Grubbs, "Procedures for Detecting Outlying Observations in Samples," *Technometrics*, vol. 11, no. 1, pp. 1–21, Feb. 1969

[3] A. C. Atkinson and D. M. Hawkins, "Identification of Outliers.," *Biometrics*, vol. 37, no. 4, p. 860, Dec. 1981

[4] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, "LOF," in *Proceedings of the 2000 ACM SIGMOD international conference on Management of data - SIGMOD '00*, 2000, pp. 93–104

[5] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection," *ACM Comput. Surv.*, vol. 41, no. 3, pp. 1–58, Jul. 2009

[6] H. Wang, M. J. Bah, and M. Hammad, "Progress in Outlier Detection Techniques: A Survey," *IEEE Access*, vol. 7, pp. 107964–108000, 2019

[7] S. Poddar and B. K. Patra, "Reduction in execution cost of k-nearest neighbor based outlier detection method," in *Communications in Computer and Information Science*, vol. 834, 2018, pp. 53–60.

[8] X. and others Ester, Martin and Kriegel, Hans-Peter and Sander, Jörg and Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise.," *Kdd*, vol. 96, pp. 226–231, 1996

[9] S. Ramaswamy, R. Rastogi, and K. Shim, "Efficient algorithms for mining outliers from large data sets," *Proc. 2000 ACM SIGMOD Int. Conf. Manag. data - SIGMOD '00*, vol. 29, no. 2, pp. 427–438, Jun. 2000

[10] J. Ha, S. Seok, and J. S. Lee, "Robust outlier detection using the instability factor," *Knowledge-Based Syst.*, vol. 63, pp. 15–23, Jun. 2014

[11] J. Yang, S. Rahardja, and P. Fränti, "Mean-shift outlier detection and filtering," *Pattern Recognit.*, vol. 115, p. 107874, Jul. 2021

[12] B. Tang and H. He, "A local density-based approach for outlier detection," *Neurocomputing*, vol. 241, pp. 171–180, Jun. 2017

[13] R. Heydari Gharaei and H. Nezamabadi-pour, "A Novel Approach to Outlier Detection Based on Inverse Neighborhood," *5th Int. Conf. Knowledge-Based Eng. Innov.*, 2019, [Online]. Available: https://civilica.com/doc/988926. (in Farsi)

[14] H.-P. Kriegel, P. Kröger, E. Schubert, and A. Zimek, "LoOP," in *Proceeding of the 18th ACM conference on Information and knowledge management - CIKM '09*, 2009, p. 1649

[15] G. O. Campos *et al.*, "On the evaluation of unsupervised outlier detection: measures, datasets, and an empirical study," *Data Min. Knowl. Discov.*, vol. 30, no. 4, pp. 891–927, Jul. 2016

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

An Ensemble Of Feature Selection Algorithms Using OWA Operator

Amin Hashemi[1], Mohammad Bagher Dowlatshahi [2]

[1] *Department of Computer Engineering, Faculty of Engineering, Lorestan University, Khorramabad, Iran,* hashemi.am@fe.lu.ac.ir

[2] *Department of Computer Engineering, Faculty of Engineering, Lorestan University, Khorramabad, Iran,* dowlatshahi.mb@lu.ac.ir

***Abstract*— This article tries to deal with the problem of ensemble feature selection as a rank aggregation procedure. As our proposed algorithm, EFS-OWA first constructs a decision matrix based on the rank of features obtained by different feature selection algorithms. To aggregate the results of different feature selection methods, we used Ordered Weighted Averaging (OWA) aggregation operator. This will allow the features that have the most satisfactory by feature selection algorithms assigned highest ranks. Some ensemble feature selection algorithms that use rank aggregation procedures are compared with EFS-OWA to prove the performance of the proposed method. Also, another comparison is conducted between EFS-OWA and basic feature selection algorithms. All these comparisons are made based on classification accuracy and the runtime of algorithms. Based on the experiment results, we can see that EFS-OWA outperforms competitive methods.**

***Index Terms* — Ensemble feature selection; Ordered weighted averaging; Rank aggregation; Decision matrix.**

I. INTRODUCTION

Nowadays, vast amounts of data are generated daily due to growing technologies like social networks and the Internet of Things (IoT). Usually, there are many high-dimensional data among these data. This dimensionality of data is due to the existence of a large number of features and training samples. Redundant and irrelevant features are included in many data that can affect the performance of learning algorithms and increase the computational complexity and learning time. Since irrelevant features cannot help the classifier predict the class labels and may mislead the learning process, redundant features are a combination of other features and cannot add any useful information. Feature selection (FS) is a powerful technique to handle mentioned problems. FS is a well-known dimensionality reduction process that removes the redundant and irrelevant features and chooses the most effective subset. FS enhances classification accuracy and decreases the computational cost, storage complexity, time complexity [1]–[5]. FS methods can be grouped based on search strategy and supervision perspectives [6], [7].

FS is categorized into three filters, wrapper, and embedded methods based on the search strategy perspective. Filters assess the features by the statistical measures and without using the classifier [8], [9]. The assessment of features is conducted during the learning in wrapper-based methods. Wrappers evaluated possible feature subsets to find the best one that provides the maximum classification accuracy among all subsets [1], [10]. At last, embedded methods, just like wrappers, use the classifier in the learning process. Unlike wrappers, it is not performing iteratively, but it is just performed once. Then the features are assessed like filter-based methods [6], [11].

According to the supervision perspective, supervised, unsupervised, and semi-supervised are three main categories of FS methods [12], [13]. Supervised FS methods consider the relationship between features and labels. These methods are used for single-label (SL) and multi-label (ML) data [14]–[16]. [12], [17]. An unsupervised method assesses the features without any knowledge of class labels [12]. Semi-supervised methods apply to datasets with few labeled instances [18].

At the same time, a new style of FS methods has been introduced, which are made from a combination of FS methods. These methods, called ensemble feature selection (EFS), combine the output of several different feature selection methods. The idea of this technique is according to the old proverb "Two heads are better than one" and that a combination of different methods can cause different aspects of the data to be evaluated from several prospective and prevent from falling into a local optimum. [19]–[21].

This paper model the ensemble FS problem to a rank aggregation procedure. At first, a decision matrix is constructed by the ranks of each feature obtained by different FS methods. Then we used the OWA as a well-known aggregation operator. To do this, we must calculate the weight of each feature selection algorithm. Thus, we consider the minimum correlation of each FS method to other methods as the weight of that algorithm in the aggregation process. Finally, the features are sorted based on the following measures and add to the feature ranking vector.

Five new EFS and four base FS algorithms are provided. All algorithms have been performed on real-world datasets to prove the performance of the EFS-OWA. The experiment shows that EFS-OWA outperforms competitive methods in terms of classification accuracy and runtime.

The structure of this paper is organized as follows: In Section II, related methods are reviewed. Section III presents the description of the fundamental concepts, and Section IV describes the OFS-OWA method in detail. Section V presents the experimental results, and in Section VI, the conclusion is provided.

II. RELATED WORKS

PEFS method [22] used a rank aggregation procedure for ensemble feature selection. In this method, the results of four rankers are combined into two vectors. These two vectors represent the average and minimum ranks for each feature. Then, these two vectors are considered as two objectives in a Pareto dominance space. At last, the features are sorted based on non-dominated sorting and crowding distance. In EFS-MCDM [10], the ensemble feature selection procedure is assumed a Multi-Criteria Decision-Making (MCDM) process. Each filter ranker is considered the criteria to evaluate the features representing the alternatives in the MCDM process. In EFS-MCDM, after the filter rankers construct a decision matrix, the VIKOR method sorts the features. In [23], some popular election methods such as Plurality voting, Borda count, and weighted Borda count have been used to ensemble the feature selection algorithms. These methods have used eight rankers in the fusion process. In reference [24], an ensemble of feature selection algorithms is proposed based on a greedy swarm optimizer using filter-based rankers for high dimensional datasets.

III. FUNDAMENTAL CONCEPTS

A. Ensemble feature selection

The process of aggregating multiple feature selection methods is called ensemble feature selection. The idea of this technique is according to the old proverb "Two heads are better than one" and that a combination of different methods can cause different aspects of the data to be evaluated from several perspectives and prevent falling into a local optimum. Machine learning tasks such as optimization problems, regression, and classification have evolved with ensemble techniques to enhance performance. Furthermore, these techniques are also effective in feature selection. There are two main groups for ensemble feature selection methods: homogeneous ensembles exploit data diversity. The data is divided into multiple partitions into these methods, and on each partition, a feature selection algorithm is executed. It should also be noted that the same feature selection algorithm must be run on all partitions. Finally, the obtained results of all partitions are aggregated to obtain the desirable feature subset. Heterogeneous ensembles focus on functions diversity. In this category, several feature selection methods are implemented on the entire data. The final result is achieved based on combining these functions' results [19], [25].

B. Ordered Weighted Averaging (OWA) operator

The Ordered Weighted Averaging (OWA) operator is a method for aggregating the results of different functions. This method was introduced by Yager [26] and is widely used in decision-making issues that there is a matter of uncertainty about them. If the decision vector for each element involves n dimension, the OWA operator maps $R^n \rightarrow R$. The weighting vector W also shows the weight of different functions that represent a value in the range $[0,1]$ to each function so that $\sum_{j=1}^{n} w_j = 1$. The OWA operator assign a score to each element using the following equation:

$$OWA(d_1, d_2, \dots, d_n) = \sum_{j=1}^{n} w_j d_j \quad (1)$$

Where d_j is the value which the j-th function assign to the element [27], [28].

IV. PROPOSED METHOD

In this section, we discuss our proposed algorithm in detail. This algorithm is a filter-based heterogeneous algorithm. In this method, we used the ensemble of four FS methods and did the aggregation process by the OWA operator. Algorithm 1 shows the step-by-step process of the proposed method.

A. Motivation

Ensemble techniques are powerful methods that can be employed to enhance the performance of FS algorithms. Every FS method may obtain a different feature subset. These various subsets of features may show local optima in the feature space. Thus, the ensemble FS techniques are employed to address this problem by performing a fusion of several FS. For this matter, we offered a rank aggregation algorithm for ensemble feature selection using the OWA operator. OWA operator is usually used in multi-criteria decision-making problems where there are conflicting criteria. In ensemble feature selection like multi-criteria decision-making problems, the final feature subset is achieved by evaluating multiple base feature selection algorithms. Thus, the OWA operator is useful for aggregating the result of different methods.

B. Proposed algorithm

In this section, the description of algorithm 1 is presented. Step 1. refers to defining an empty vector (F) to store the final feature subset.

Algorithm 1: An ensemble of feature selection algorithms using OWA operator (EFS-OWA)

Input: $N \times 1$ label vector $\boldsymbol{Y}$, $N \times M$ Feature data matrix $\boldsymbol{X}$
Output: vector F *representing the ranking of the features*
1. $\boldsymbol{F} = \emptyset$;
2. **R1 = The rankings achieved by CFS method**
3. **R2 = The rankings achieved by Fisher − Score method**
4. **R3 = FThe rankings achieved by LLCFS method**
5. **R4 = The rankings achieved by MIC method**
6. P=[R1,R2,R3,R4];
7. [**m**, **n**] = **size**(**P**)
8. Construct a decision matrix (D)
9. for i=1:n
10. for j=1:m
11. T=P(j,i)
12. D(T, i)= j
13. end
14. end
15. **for** i=1:n
16. **for** j=1:n
17. Correlation (i,j)=Corr (D(:,i),D(:,j))
18. **end**
19. **end**
20. W=min (Correlation)
21. W=W./(sum(W))
22. **for** i=1:m
23. OWA(i)=sum(D(i,:).*W)
24. **end**
25. F = sort the features in ascending order based on their values in OWA;

Steps 2 to 5 indicates the implementation of four FS methods (Fisher-Score [29], MIC [29], LLCFS [30], and CFS [31]) and saving their outputs (R1 to R4) as rank vectors. In this method, we have decided to use a combination of relevancy-based and redundancy-based FS methods to stable the convergence and diversity of the algorithm. For this matter, Fisher-score and MIC methods represent our relevancy-based metrics, and the LLCFS and CFS methods are our redundancy-based methods. The feature ranking vectors (R1 to R4) are containing the rankings of the following methods. It means that the first value of these vectors shows the highest rank feature index. The construction of the following vectors is shown as follows:

$$\mathrm{R1} = \{\mathrm{r}_{1,1}, \mathrm{r}_{1,2}, \dots\dots, \mathrm{r}_{1,\mathrm{M}}\}\ \#\mathrm{CFS} \quad (2)$$

$$\mathrm{R2} = \{\mathrm{r}_{2,1}, \mathrm{r}_{2,2}, \dots\dots, \mathrm{r}_{2,\mathrm{M}}\}\#\mathrm{Fisher-score} \quad (3)$$

$$\mathrm{R3} = \{\mathrm{r}_{3,1}, \mathrm{r}_{3,2}, \dots\dots, \mathrm{r}_{1,\mathrm{M}}\}\#\mathrm{LLCFS} \quad (4)$$

$$\mathrm{R4} = \{\mathrm{r}_{4,1}, \mathrm{r}_{4,2}, \dots\dots, \mathrm{r}_{4,\mathrm{M}}\}\#\mathrm{MIC} \quad (5)$$

Step 6 presents the process of constructing the ranking matrix (P). Every column of the matrix P indicates the ranking vectors (R1 to R4), and each row refers to the features. Step 7 of the algorithm shows the calculation of the dimensions of the matrix P, in which m refers to the rows (features), and n shows the number of columns (FS methods). Matrix P is structured as follows:

$$\mathrm{P} = \begin{bmatrix} r_{11} & r_{21} & r_{31} & r_{41} \\ r_{12} & r_{22} & r_{32} & r_{42} \\ \vdots & \vdots & \vdots & \vdots \\ r_{1M} & r_{2M} & r_{3M} & r_{4M} \end{bmatrix} \quad (6)$$

Steps 8 to 14 shows a process that each feature is assigned to its corresponding rank based on the matrix P to build the decision matrix (D):

$$\mathrm{D} = \begin{bmatrix} d_{11} & d_{21} & d_{31} & d_{41} \\ d_{12} & d_{22} & d_{32} & d_{42} \\ \vdots & \vdots & \vdots & \vdots \\ d_{1M} & d_{2M} & d_{3M} & d_{4M} \end{bmatrix} \quad (7)$$

In steps 15 to 19, we calculated the minimum correlation of each method against the others. The correlation of two random FS methods is defined as follows:

$$Corr(i,j) = Cov(D(:,i), D(:,j))/var(D(:,i))var(D(:,j)) \quad (8)$$

The obtained matrix is:

$$\mathrm{Correlation} = \begin{bmatrix} Corr(1,1) & Corr(1,2) & Corr(1,3) & Corr(1,4) \\ Corr(2,1) & Corr(2,2) & Corr(2,3) & Corr(2,4) \\ Corr(3,1) & Corr(3,2) & Corr(3,3) & Corr(3,4) \\ Corr(4,1) & Corr(4,2) & Corr(4,3) & Corr(4,4) \end{bmatrix} \quad (9)$$

In step 20, the minimum correlation value of each feature selection method is considered the weight of that method in the aggregation process. This will let the method that has the least similarity with other methods assigns a higher weight. The obtained vector is:

$$W = \begin{bmatrix} \min(correlation(1,:) \\ \min(correlation(2,:) \\ \min(correlation(3,:) \\ \min(correlation(4,:) \end{bmatrix} \quad (10)$$

Because $\sum_{j=1}^{n} W_j = 1$, we normalized the weights by dividing each weight by the sum of weights:

$$W = \begin{bmatrix} \mathrm{W}_1/sum(W) \\ \mathrm{W}_2/sum(W) \\ \mathrm{W}_3/sum(W) \\ \mathrm{W}_4/sum(W) \end{bmatrix} \quad (11)$$

In steps 22 to 24, we have calculated the OWA score for each feature based on the matrix D and vector W using Eq. 1.

Finally, the features are sorted according to their corresponding value in OWA in ascending order. The results are stored in the vector F, in which the users can select a desired number of features.

V. EXPERIMENTAL STUDIES

The performance of the proposed is evaluated based on a two-phase comparison: In the first phase, four simple FS methods are considered: MIC [29], Fisher-Score [29], CFS [31], and LLCFS [30]. In the second phase, the comparison with ensemble FS methods is performed. These methods are E-Borda [23] which uses the Borda-count method, E-WBorda [23] uses weighted Borda-count, E-Plu[23] uses plurality voting, and EFSPF [32], and EFS-MCDM [10] that use rank aggregation procedures.

A. Datasets

Five real-world datasets are considered to assess the performance of the proposed method and competitive methods in our two-phase comparison. The description of these datasets is presented in Table I.

TABLE I. DESCRIPTION OF DATASETS

Dataset	Instances	Features	Type
Pomeroy	60	7129	Biological
Lymphoma	96	4027	Microarray Data
Arcene	200	10001	Mass Spectrometry
ORL	400	1025	Image
Lung	203	3313	Biological

B. Performance evaluation criteria

Accuracy [25] measure is considered for measuring the classification performance of the EFS-OWA and comparing methods. Accuracy metric is defined based on the following concepts:

1) False-positive (FP): percentage of negative examples incorrectly classified as positive.

2) True positive (TP): percentage of positive examples correctly classified as so.

3) False-negative (FN): percentage of positive examples incorrectly classified as negative.

4) True negative (TN): percentage of negative examples correctly classified as so.

The accuracy measure is defined as follows:

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \qquad (12)$$

C. Results

The recommended FS methods for the ensemble process based on the corresponded paper are considered in all competitive methods. The evaluation of classification accuracy for all comparing methods is performed based on the K-nearest neighbors (K-NN) classifier. The number of neighbors in the following classifier is set to 5 in all experiments.

We have set 60% of the instances in each dataset as training data randomly and the remaining 40% as the test data. 30 separate runs obtain the reported result of each method on every dataset, and the presented values are the average value of the following 30 runs, where *m* features are selected as representative from the features extracted from each algorithm, and the test is performed on them. The number of representatives is between 10 and 100 and a factor of 10 ($m \in \{10, 20, 30, 40, 50, 60, 70, 80, 90, 100\}$).

Figures. 1 to 5 represent the classification accuracy of the proposed methods and simple FS methods as our phase one comparison. The horizontal axis indicates the number of selected features in these figures, and the vertical axis shows the comparison criterion.

For phase two comparison, Tables II to VI shows the classification performance for accuracy criterion comparing ensemble FS methods. Also, the algorithm runtimes are reported in Table VII.

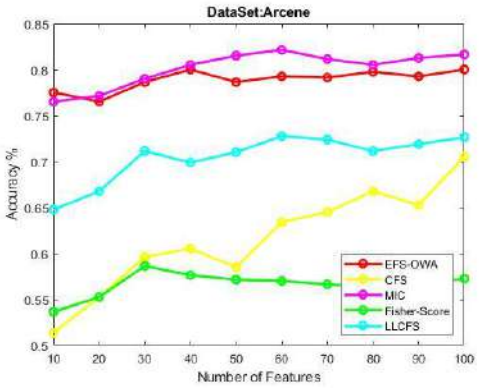


Figure 1. Accuracy for Arcene dataset

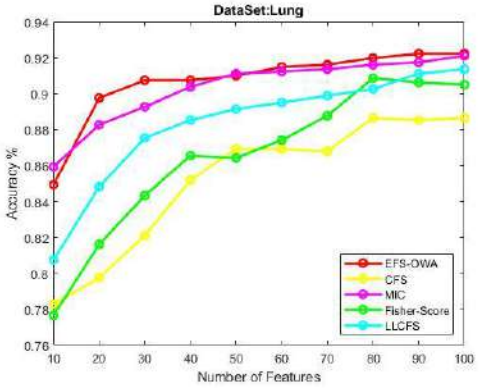


Figure 2. Accuracy for Lung dataset

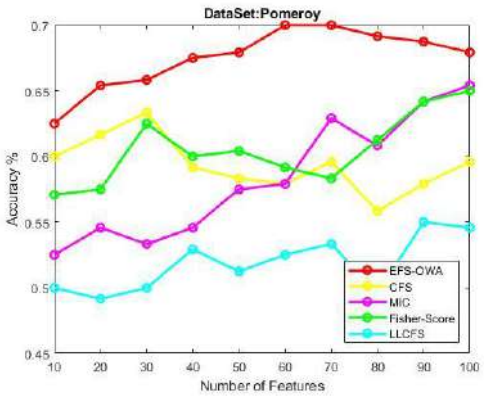


Figure 3. Accuracy for Pomeroy dataset

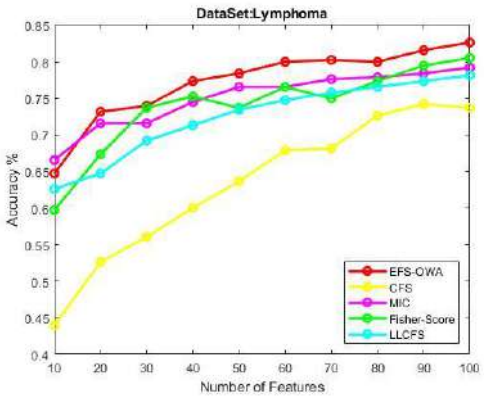


Figure 4. Accuracy for Lymphoma dataset

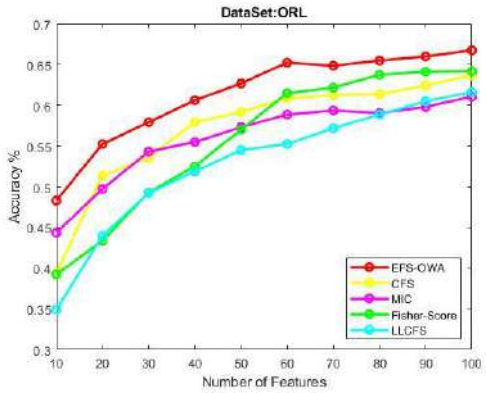


Figure 5. Accuracy for ORL dataset

TABLE II. ACCURACY OF ENSEMBLE METHODS FOR LUNG DATASET

M/fea	EFS-OWA	E-Borda	E-Wborda	E-Plu	EFSPF	EFS-MCDM
10	**0.8494**	0.8148	0.8407	0.8185	0.7963	0.8111
20	**0.8975**	0.8420	0.8605	0.8704	0.8395	0.8642
30	**0.9074**	0.8667	0.8580	0.8778	0.8481	0.8926
40	**0.9074**	0.8753	0.8790	0.8815	0.8691	0.8951
50	**0.9099**	0.8926	0.8778	0.8815	0.8778	0.8914
60	**0.9148**	0.8914	0.8926	0.8988	0.8840	0.9062
70	**0.9160**	0.8963	0.9086	0.9000	0.8938	0.9123
80	**0.9198**	0.8988	0.9160	0.9062	0.8951	0.9099
90	**0.9222**	0.9074	0.9160	0.9074	0.9099	0.9198
100	0.9222	0.9074	**0.9247**	0.9074	0.9185	0.9099

TABLE III. ACCURACY OF ENSEMBLE METHODS FOR ARCENE DATASET

M/fea	EFS-OWA	E-Borda	E-Wborda	E-Plu	EFSPF	EFS-MCDM
10	**0.7750**	0.6637	0.4662	0.6900	0.3349	0.4700
20	**0.7650**	0.7087	0.4662	0.7300	0.4275	0.5356
30	**0.7863**	0.7137	0.4662	0.7488	0.4794	0.5688
40	**0.8000**	0.7113	0.4662	0.7550	0.5100	0.6031
50	**0.7863**	0.7438	0.4662	0.7537	0.5100	0.6000
60	**0.7925**	0.7375	0.4875	0.7762	0.5350	0.6219
70	**0.7913**	0.7388	0.5212	0.7787	0.5569	0.6144
80	**0.7975**	0.7337	0.6175	0.7775	0.5662	0.6169
90	**0.7925**	0.7512	0.6850	0.7913	0.5694	0.6288
100	**0.8000**	0.7475	0.7250	0.7788	0.5750	0.6275

TABLE IV. ACCURACY OF ENSEMBLE METHODS FOR POMEROY DATASET

M/fea	EFS-OWA	E-Borda	E-Wborda	E-Plu	EFSPF	EFS-MCDM
10	0.6250	0.3969	0.4269	0.3688	**0.6417**	0.6167
20	**0.6542**	0.4963	0.5069	0.4644	0.5917	0.6323
30	**0.6583**	0.5312	0.5550	0.4794	0.5917	0.5958
40	**0.6750**	0.5569	0.5887	0.4994	0.6083	0.6000
50	**0.6792**	0.5819	0.6144	0.5219	0.5625	0.5875
60	**0.7000**	0.6012	0.6169	0.5475	0.5708	0.6042
70	**0.7000**	0.6131	0.6294	0.5575	0.5667	0.5917
80	**0.6917**	0.6169	0.6263	0.5619	0.5958	0.6042
90	**0.6875**	0.6269	0.6425	0.5575	0.6000	0.6167
100	**0.6792**	0.6319	0.6438	0.5631	0.6000	0.6063

TABLE V. ACCURACY OF ENSEMBLE METHODS FOR LYMPHOMA DATASET

M/fea	EFS-OWA	E-Borda	E-Wborda	E-Plu	EFSPF	EFS-MCDM
10	**0.6474**	0.5150	0.5050	0.4550	0.5974	0.6167
20	**0.7316**	0.5850	0.5200	0.4800	0.6579	0.6333
30	**0.7395**	0.5750	0.5450	0.4900	0.6974	0.5958
40	**0.7737**	0.5800	0.5500	0.5400	0.7053	0.6000
50	**0.7842**	0.6250	0.5650	0.5400	0.7053	0.5875
60	**0.8000**	0.6300	0.5800	0.5400	0.7237	0.6042
70	**0.8026**	0.6350	0.5850	0.5700	0.7316	0.5917
80	**0.8000**	0.6550	0.5950	0.5950	0.7421	0.6042
90	**0.8158**	0.6250	0.5850	0.5700	0.7632	0.6167
100	**0.8263**	0.6350	0.6200	0.6150	0.7605	0.6083

TABLE VI. ACCURACY OF ENSEMBLE METHODS FOR ORL DATASET

M/fea	EFS-OWA	E-Borda	E-Wborda	E-Plu	EFSPF	EFS-MCDM
10	**0.4831**	0.3969	0.4269	0.3688	0.3394	0.4700
20	**0.5525**	0.4963	0.5069	0.4644	0.4275	0.5356
30	**0.5794**	0.5312	0.5550	0.4794	0.4794	0.5688
40	**0.6063**	0.5569	0.5887	0.4994	0.5100	0.6031
50	**0.6269**	0.5819	0.6144	0.5219	0.5100	0.6000
60	**0.6525**	0.6012	0.6169	0.5475	0.5350	0.6219
70	**0.6488**	0.6131	0.6294	0.5575	0.5569	0.6144
80	**0.6550**	0.6169	0.6263	0.5619	0.5662	0.6169
90	**0.6600**	0.6269	0.6425	0.5575	0.5694	0.6288
100	**0.6675**	0.6319	0.6438	0.5631	0.5750	0.6275

TABLE VII. THE AVERAGE RUNTIME OF EFS-OWA AGAINST THE ENSEMBLE FS METHODS

M/fea	EFS-OWA	E-Borda	E-Wborda	E-Plu	EFSPF	EFS-MCDM
Arcene	**25.74**	1220.22	1110.07	1140.27	16.70	26.37
Pomeroy	**4.62**	1670.42	1791.01	1688	501.83	4.89
Lymphoma	3.41	1670.02	1791.81	1688	114.42	**3.02**
ORL	**25.37**	2008.31	2013.24	2010	16.70	24.01
Lung	14.60	3059.25	3063.93	3057	71.18	**2.43**

D. Discussion

EFS-OWA method is a new algorithm for ensemble feature selection that uses the OWA operator to combine the simple FS methods based on a rank aggregation strategy. In this method, we have used a combination of redundancy-based and relevancy-based FS methods to balance the diversity and convergence of the algorithm. This process causes the selection of less redundant and most relevant feature subset. We have used the MIC and Fisher-score methods as our relevancy-based method. Also, we have chosen CFS and LLCFS methods as our redundancy-based metrics. Thus, the constructed decision matrix is a fusion of relevancy and redundancy equally. The reported results indicate the superiority of the EFS-OWA method against competitive methods. Based on reported results in Table VII, the proposed method is performed in a reasonable time. In three cases of all performing datasets, EFS-OWA has the shortest runtime.

VI. CONCLUSION

This article proposes an ensemble feature selection method called the EFS-OWA, which models the features selection procedure to a rank aggregation process. EFS-OWA is a filter-based method and uses a heterogeneous ensemble strategy. In this paper, first, we calculated the ranks of features based on several FS methods and constructed a decision matrix. The

decision matrix is delivered to the OWA operator to perform the ensemble process in the next step. The results of different datasets indicate the efficiency and optimality of the proposed method since the convergence and diversity of the algorithm are balanced by the variety of feature selection methods. Also, due to the simple calculation in the process of the algorithm, the method performs in a short runtime. We intend to use ensemble techniques in other feature selection applications for our feature works and combine them with various optimization tasks in machine learning.

ACKNOWLEDGMENT

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors

REFERENCES

[1] J. Miao and L. Niu, "A Survey on Feature Selection," in *Procedia Computer Science*, 2016, vol. 91, pp. 919–926, doi: 10.1016/j.procs.2016.07.111.

[2] W. C. Mlambo, N, "A survey and comparative study of filter and wrapper feature selection techniques," *Int. J. Eng. Sci.*, vol. 5, no. 8, pp. 57–67, 2016, [Online]. Available: https://pdfs.semanticscholar.org/fe2c/a244dd5473270ce0a8e2a387349b9831cf47.pdf.

[3] J. Cai, J. Luo, S. Wang, and S. Yang, "Feature selection in machine learning: A new perspective," *Neurocomputing*, vol. 300, pp. 70–79, 2018, doi: 10.1016/j.neucom.2017.11.077.

[4] H. Bayati, M. B. Dowlatshahi, and M. Paniri, "MLPSO: A Filter Multi-label Feature Selection Based on Particle Swarm Optimization," in *2020 25th International Computer Conference, Computer Society of Iran (CSICC)*, Jan. 2020, pp. 1–6, doi: 10.1109/CSICC49403.2020.9050087.

[5] M. B. Dowlatshahi, V. Derhami, and H. Nezamabadi-pour, "Fuzzy particle swarm optimization with nearest-better neighborhood for multimodal optimization," *Iran. J. Fuzzy Syst.*, 2019, doi: 10.22111/ijfs.2019.5054.

[6] J. Li *et al.*, "Feature Selection: A Data Perspective," *ACM Comput. Surv.*, vol. 50, 2017, doi: 10.1145/3136625.

[7] M. Paniri, M. B. Dowlatshahi, and H. Nezamabadi-pour, "Ant-TD: Ant colony optimization plus temporal difference reinforcement learning for multi-label feature selection," *Swarm Evol. Comput.*, vol. 64, p. 100892, Jul. 2021, doi: 10.1016/j.swevo.2021.100892.

[8] B. Venkatesh and J. Anuradha, "A review of Feature Selection and its methods," *Cybern. Inf. Technol.*, vol. 19, no. 1, pp. 3–26, 2019, doi: 10.2478/CAIT-2019-0001.

[9] A. Hashemi, M. Bagher Dowlatshahi, and H. Nezamabadi-pour, "VMFS: A VIKOR-based multi-target feature selection," *Expert Syst. Appl.*, p. 115224, May 2021, doi: 10.1016/j.eswa.2021.115224.

[10] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-pour, "Ensemble of feature selection algorithms: a multi-criteria decision-making approach," *Int. J. Mach. Learn. Cybern.*, pp. 1–21, Jun. 2021, doi: 10.1007/s13042-021-01347-z.

[11] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-pour, "MFS-MCDM: Multi-label feature selection using multi-criteria decision making," *Knowledge-Based Syst.*, p. 106365, Aug. 2020, doi: 10.1016/j.knosys.2020.106365.

[12] R. Zhang, F. Nie, X. Li, and X. Wei, "Feature selection with multi-view data: A survey," *Inf. Fusion*, vol. 50, pp. 158–167, 2019, doi: 10.1016/j.inffus.2018.11.019.

[13] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-Pour, "A bipartite matching-based feature selection for multi-label learning," *Int. J. Mach. Learn. Cybern.*, vol. 12, no. 2, pp. 459–475, Feb. 2021, doi: 10.1007/s13042-020-01180-w.

[14] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-pour, "MGFS: A multi-label graph-based feature selection algorithm via PageRank centrality," *Expert Syst. Appl.*, vol. 142, p. 113024, 2020, doi: 10.1016/j.eswa.2019.113024.

[15] A. Hashemi and M. B. Dowlatshahi, "MLCR: A Fast Multi-label Feature Selection Method Based on K-means and L2-norm," in *2020 25th International Computer Conference, Computer Society of Iran (CSICC)*, Jan. 2020, pp. 1–7, doi: 10.1109/CSICC49403.2020.9050104.

[16] A. Hashemi, M. Bagher Dowlatshahi, and H. Nezamabadi-pour, "An efficient Pareto-based feature selection algorithm for multi-label classification," *Inf. Sci. (Ny).*, vol. 581, pp. 428–447, Dec. 2021, doi: 10.1016/j.ins.2021.09.052.

[17] R. B. Pereira, A. Plastino, B. Zadrozny, and L. H. C. Merschmann, "Categorizing feature selection methods for multi-label classification," *Artif. Intell. Rev.*, vol. 49, no. 1, pp. 57–78, Jan. 2018, doi: 10.1007/s10462-016-9516-4.

[18] R. Sheikhpour, M. A. Sarram, S. Gharaghani, and M. A. Z. Chahooki, "A Survey on semi-supervised feature selection methods," *Pattern Recognit.*, vol. 64, pp. 141–158, 2017, doi: 10.1016/j.patcog.2016.11.003.

[19] V. Bolón-Canedo and A. Alonso-Betanzos, "Ensembles for feature selection: A review and future trends," *Inf. Fusion*, vol. 52, pp. 1–12, 2019, doi: 10.1016/j.inffus.2018.11.008.

[20] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-Pour, "Minimum redundancy maximum relevance ensemble feature selection: A bi-objective Pareto-based approach," *J. Soft Comput. Inf. Technol.*, 2021.

[21] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-pour, "Gravitational Search Algorithm," in *Handbook of AI-based Metaheuristics*, 2021, p. 32.

[22] A. Hashemi, M. Bagher Dowlatshahi, and H. Nezamabadi-pour, "A pareto-based ensemble of feature selection algorithms," *Expert Syst. Appl.*, vol. 180, p. 115130, Oct. 2021, doi: 10.1016/j.eswa.2021.115130.

[23] P. Drotár, M. Gazda, and L. Vokorokos, "Ensemble feature selection using election methods and ranker clustering," *Inf. Sci. (Ny).*, vol. 480, pp. 365–380, 2019, doi: 10.1016/j.ins.2018.12.033.

[24] M. B. Dowlatshahi, V. Derhami, and H. Nezamabadi-Pour, "Ensemble of filter-based rankers to guide an epsilon-greedy swarm optimizer for high-dimensional feature subset selection," *Inf.*, vol. 8, no. 4, 2017, doi: 10.3390/info8040152.

[25] V. Bolón-Canedo and A. Alonso-Betanzos, "Evaluation of ensembles for feature selection," in *Intelligent Systems Reference Library*, vol. 147, 2018, pp. 97–113.

[26] R. R. Yager, "On Ordered Weighted Averaging Aggregation Operators in Multicriteria Decisionmaking," *IEEE Trans. Syst. Man Cybern.*, vol. 18, no. 1, pp. 183–190, 1988, doi: 10.1109/21.87068.

[27] J. M. Merigó, D. Palacios-Marqués, and P. Soto-Acosta, "Distance measures, weighted averages, OWA operators and Bonferroni means," *Appl. Soft Comput. J.*, vol. 50, pp. 356–366, 2017, doi: 10.1016/j.asoc.2016.11.024.

[28] R. R. Yager, "OWA aggregation of multi-criteria with mixed uncertain satisfactions," *Inf. Sci. (Ny).*, vol. 417, pp. 88–95, 2017, doi: 10.1016/j.ins.2017.06.037.

[29] R. O. Duda, P. E. Hart, and D. G. Stork, "Pattern classification," *New York John Wiley, Sect.*, vol. 10, p. l, 2001.

[30] H. Zeng and Y. M. Cheung, "Feature selection and kernel learning for local learning-based clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 8, pp. 1532–1547, 2011, doi: 10.1109/TPAMI.2010.215.

[31] K. Michalak and H. Kwasnicka, "Correlation based feature selection method," *Int. J. Bio-Inspired Comput.*, vol. 2, no. 5, pp. 319–332, 2010, doi: 10.1504/IJBIC.2010.036158.

[32] M. Joodaki, M. B. Dowlatshahi, and N. Z. Joodaki, "An ensemble feature selection algorithm based on PageRank centrality and fuzzy logic," *Knowledge-Based Syst.*, vol. 233, p. 107538, Dec. 2021, doi: 10.1016/j.knosys.2021.107538.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

An Ensembled Text-based Multi-Stacked-CNN-BILSTM for Rumor Detection on Social Media

1st Rashid Behzadidoost, 2nd Habib Izadkhah

1st Department of Computer Science, *University of Tabriz*, Tabriz, Iran, r.behzadi@tabrizu.ac.ir
2nd Department of Computer Science, *University of Tabriz*, Tabriz, Iran, izadkhah@tabrizu.ac.ir

***Abstract*— Social networks heave became a part of people's life. Although there are many positive sides like eCommerce and easy communication through social networks, the existence of rumors is the negative side of social networks. With the least effort, a rumor can be quick to many users disseminated leading to unpleasant social phenomena. In the literature, there are some machine learning and deep learning approaches to automatically detect rumors. Most of the previous works fail to detect the rumors in high performance. In this paper, we present an ensembled multi-stacked-CNN-BILSTM model to detect rumors on social media. This deep learning model just considers the text of users and detects that whether a text message is a rumor or not. In the proposed model, each multi-stacked-CNN-BILSTM makes a prediction, and the label that voted maximum as predicted label is detected. We have evaluated the proposed model on the real public data set of PHEME. The proposed model achieved accuracy improvement over the ten state-of-the-art machine learning and deep learning models.**

***Index Terms*— Rumor Detection, Deep Learning, Twitter, Natural Language Processing**

I. Introduction

Nowadays, Twitter is one of the most popular social networks used by users. Some Twitter users use Twitter to spread their opinions and others to keep abreast of the latest global events [1]. Millions of text messages are tweeted daily on Twitter. Hence, Twitter has a lot of up-to-date big data. With this story, Twitter has become one of the best resources for data scientists and researchers for natural language processing[1] [1].

Despite the many benefits of social networks [2], one of its disadvantages is the existence of fake news and rumors. Twitter is no exception to this rule, and it spreads rumors and fake news on various topics. In social networks, a rumor is a text message that is widely disseminated among users and aims at the widespread destruction of an entity, large corporations, and so on. Fake news is not well-organized and is not spread as speedily as a rumor. As social networks have become a part of people's lives these days, fake news and rumors can act a very destructive role in people's lives and businesses. For instance, in the coronavirus epidemic, rumors encourage people to be not vaccinated or encourage them to consume chemicals detergents to prevent getting infected with the coronavirus. To prevent and minimize such damages, there is a necessity for models that automatically and efficiently detect the rumors. In this study, we present a deep learning model that can distinguish rumors from non-rumors. Generally, the feature extraction for rumor detection is either content-based or context-based [3]. Content features extensively are used in NLP [3]. Content features are extracted from the types of lexical-based, syntactic-bases, and semantic-based [3]. Syntactic-based refer to nouns, verbs, adjectives, and POS patterns [3]. Lexical-based consider the words of a text, and it is the straightest way for the natural language processing [3]. Semantic-based extracts the features from NLP fields such as sentiment analysis, opinion mining, and topic detection [3]. As the words of the posts directly are used to spread a rumor, the content features are the most intuitive way to investigate the rumor/non-rumor structure. However, the word set of rumor and non-rumor may be very similar, and it is to be very hard to distinguish between them.

Context features are extracted by considering appropriate information concerning the actual posts [3]. Mainly, context features have consisted of types user-based and network-based [3]. In user-based, the pieces of information such as the number of followers, country, friends, and the number words of a post are considered, while in network-based, propagation structure and graphs are used.

Context features are relevant information of a post and are directly not pointed to rumor structure, while the content features (especially lexical features) are directly pointed. In lexical-based, there are many intuitive features such as writing style, negation words, capitalized words, phrases, and swear words. Hence, in this study, we intend to utilize lexical features for rumor detection [4]. ensemble learning is a technique

[1]NLP

that combines multiple learning models to make a decision; actually, ensembles are based on the wisdom of the crowd. There is much research done for the ensemble of machine learning models [4]. Recently, the ensemble of deep learning models is becoming an active area of research [5]. Works like [6] [7] have shown that ensemble deep learning models outperforms. In this study, we present a maximum-voting ensembled multi-stacked-CNN-BILSTM for rumor detection. The main contributions of this paper can be summarized as:

1) Proposing a multi-stacked-CNN-BILSTM for rumor detection. Rumor spreaders mix words in a deceitful way that may experts cannot recognize. Hence, employing a model to extract discriminative features to distinguish between rumors and non-rumors is necessary. The proposed multi-stacked-CNN-BILSTM can extract discriminative features that the reported accuracy for this model is 86.80%.
2) Using the maximum-voting ensemble for the proposed multi-stacked-CNN-BILSTM. For being robust in rumor detection, we ensembled the multi-stacked-CNN-BILSTM with its variations, where the results demonstrated that the ensembled model outperformed the non-ensembled models.

The hypotheses of this paper are as follows: (1) Employing a multi-stacked-CNN-BILSTM in a convolutional way for extracting global discriminative features improves the performance. (2) Ensembling a deep learning model with its variations reduce the dependency of a deep learning model with its hyperparameters and improves the performance.

II. Related works

Rumor detection is an active area of research. Ma et al. [8] employed a recurrent neural networks model to detect rumors. In their work, they have shown that their method outperforms on the dataset of Twitter and Weibo, respectively. Another work in the area of rumor detection was for the work of Kumar et al. [9]. In their empirical evaluations, they noted that the best accuracy in their work on the PHEME dataset was 78.40%. Ruchansky et al. [10] presented a hybrid deep model for fake news detection, where their model used combinational features. The next model, for rumor detection was a long short-term memory (LSTM) classifier [11]. The authors in [11], before classifying performed a hashtag normalization and hashtag clustering. In contrast to previous works, Nguyen [12] focused on the detection time of a rumor. Xu et al. in 2020 [13], employed a topic-driven rumor detection model, where utilized the topic of the given tweets for rumor detection. They achieved an accuracy of 82.66% for the PHEME data set compared the method with 11 studied methods. In another work, WU et al. [14] using propagation graph neural network with attention mechanism employed a four-class classification in PHEME data set. Zhang et al. [15] employed a lightweight propagation path aggregating network with the neural topic, for the PHEME data set, they achieved an accuracy of 80%. In another work, for the PHEME data set, Ajao et al. [16] using an LSTM-CNN achieved an accuracy of 90.38%.

There are some relevant works for ensemble deep learning. Zheng et al. [17] classified the EEG signals using ensemble deep learning. Using bagging theory with LSTMs, they joined the weak classifiers with their weight distribution and employed a voting policy to predict the labels. In another work, for detecting covid19, Zhou et al. [7] employed a maximum-voting ensemble with AlexNet Softmax, GoogleNet Softmax, and ResNet Softmax. Lee et al. [18] presented ensembled LSTM, where ensembled the short term, medium-term and long term memories for action recognition. In another study, by using ensemble concepts, the authors in [19] integrated the obtained weights of the Convolution Residual Network, Deep Belief Network, and Deep Auto Encoder.

III. The proposed model

In this section, the components of the proposed model are described. Machine learning models with complex problems do not perform as well as deep learning. In machine learning, the features are extracted by experts, while in deep learning, the features automatically are extracted. In this respect, feature extraction is not efficient in terms of time because it takes hours to perform. In some cases, experts even cannot detect the proper features. With this justification, we intend to employ deep learning models to detect rumors. Rumor detection in this study is a text classification task. Text classification tasks are high-dimensional problems, and development and performance outperform if the features to be extracted automatically.

a) Word Embedding: Most machine learning and deep learning models to interpret words need to transform them into a numerical representation. Word embedding is a representation of words into the continuous vector spaces, where semantic relation between words is considered, and words that have similar meaning have similar representation. One-hot encoding is one of the simplest form representations suffering from high dimensional spaces, larges sparse matrix, and no embedding for semantic information. Word embedding approaches address the limitations of one-hot encoding. To map each word to a vector, there are pre-trained and fixed embeddings. The well-known pre-trained embeddings are Word2Vec, and Glove that can capture the semantic information among the words (i.e. the more similar words the more similar representation). In fixed embedding, the words are just mapped to the numerical vector spaces.

b) Convolutional Neural Networks (CNN): CNNs are a feed-forward network that are utilized to extract informative local patterns in large sequences of various sizes. CNN employs a non-linear learned function to a window of size k. In each time step, for reaching the total sequence, the window slides. As a result, it returns a value that represents the data from that time window. From employing a filter to each window, an n-dimensional vector is obtained. By employing a pooling operation, the obtained vectors are fused into a single vector [20]. The vector represents a view of the whole sequence [20]. Unfortunately, CNNs can't take a global pattern of large sequences.

c) Data: For all evaluations, we considered the well-known data set of PHEME. This data set has 5802 tweets that 3830 and 1972 are non-rumor and rumors, respectively.

d) Long Short Term Memory: As CNNs are the fee-forward neural network, they can not capture the dependency among the tokens of a sequence. Recurrent neural networks (RNN) to address the issue are proposed. In RNN, by some internal gate, the previous inputs and states are memorized. By Using a recurrent neural network, backpropagation error [21] signals flowing backward may vanished. It is so-called the vanishing gradient. LSTMs are a solution to address the vanishing gradient. The LSTMs can save the error flow constant by gates of input and forget [21]. These gates determine how much data can be saved.

The input gate i_t checks the input x_t, while the forget gate f_t controls the output of the previous unit y_{t-1} [20]. eq:1 and eq:2 indicate the calculations of i_t and f_t. c_t save the information of time steps and k_t in eq:3 is for hidden layer. For i_t, W_i, U_i, and b_i are the weight matrices and the bias and for f_t, U_f, W_f and b_f are the weight matrices and bias, and for k_t, U_k, W_k and b_k are the weight matrices and bias [20].

eq:4 indicates the calculation of next c_t and eq:5 indicates the calculation of output; Similarly, U_y, W_y and V_y, are weight matrix and bias vector.

$$i_t = activation(dot(y_{t-1}, U_i) + dot(x_t, W_i) + b_i) \quad (1)$$

$$f_t = activation(dot(y_{t-1}, U_f) + dot(x_t, W_f) + b_f) \quad (2)$$

$$k_t = activation(dot(y_{t-1}, U_k) + dot(x_t, W_k) + b_k) \quad (3)$$

$$c_{t+1} = i_t \times k_t + c_t \times f_t \quad (4)$$

$$y_t = activation(dot(y_{t-1}, U_y) + dot(x_t, W_y)) + dot(y_t, V_y) + b_y \quad (5)$$

e) Bidirectional LSTMs: Bidirectional LSTMs are a type of neural network. In contrast to LSTMs that just consider one direction to train, BILSTMs will consider not only the forward direction but also the backward direction.

f) Proposed model: To distinguish between structures of the rumors and non-rumors, local features and the dependency among them are considered. To extract the pattern of the local features, it is employed CNNs with different kernels[2] and max-poll size as [22] and to extract many different features a stacked-CNN is employed. To extract the discriminative features, the dependency of local features by a BILSTM is captured. As text classification problems are high dimensional, extracting many proper features increase the performance. Hence, we considered multiple stacked-CNN-BILSTM.

The proposed deep learning model is depicted in fig:1.

At first, in the input layer block, the tweets from the dataset are extracted, and these textual data by Glove to the numeric vectors are transformed. In the 3rd step, CNN-BILSTM layers are created, where four stacked-CNN with different hyperparameters and four BILSTMs with the same Parameter are considered. For each stacked-CNN, a different number of CNN is considered, and each one has its different kernel and max-pooling. From each CNN in the stack, to reduce the number of parameters of each CNN, a max-pooling is taken.

[2]A kernel is a filter that is to be to pass over the input

TABLE I
HYPERPARAMETERS FOR CNNS

	Stacked CNN1	Stacked CNN2	Stacked CNN3	Stacked CNN4
Filter	128	256	64	32
Kernel	2,3,4,5,7,9,20	2,3,4,5,6,7,8	2,3,4,5,7,9	3,2,4,5,6,8,7
Max-pool1	2,3,4,5,7,9,20	2,3,4,5,6,7,8	2,3,4,5,7,9	3,2,4,5,6,8,7
Max-pool2	96	128	20	10

In this study, we have considered four stacked-CNNs that stacked CNN1, CNN2, and CNN4 have seven CNNs, and the stacked-CNN3 has 6 CNNs. The CNNs in each stack are concated, and from concated CNNs, a max-pooling is taken. In the next step, the output of each Stacked CNN is fed into a BILSTM, where the considered BILSTMs have the same parameters. In the 4th step, applying the matrix addition for each pair of the matrix output of stacked-CNN-BILSTM is performed. In step 5, from the result of the matrix additions in step 4, an average is computed. Taking average from the set of matrices gives an equal distribution of the data. In step 6, this averaged matrix is normalized. This normalization distributes the data in the center with approximately 0 with a standard deviation of 1. The normalization by eq:6 is computed. η is the output of step 5, ρ and ν are mean and variance of η. It is employed two fully connected layers that utilize feature extraction for rumor detection. In step 7, a fully connected layer with 50 nodes and an activation function of relu, and in step 8, a fully connected layer with two nodes and an activation function softmax are utilized.

$$Normalization(\eta, \rho, \nu) = \frac{\eta - \rho}{sqrt(\nu)} \quad (6)$$

g) Ensembled multi-stacked-CNN-BILSTM: For being robust on prediction, we have considered the maximum-voting ensembled approach for multi-stacked-CNN-BILSTM. In a simple statement, the proposed model using the wisdom of the crowd performs the predictions. To construct the models, we just changed the hyperparameters of multi-stacked-CNN-BILSTM. Deep learning models have hyperparameters such as the number of hidden layers, maximum sequence length, type of embedding, activation function, and the number of BILSTM nodes. These hyperparameters can impact the overall performance. One of the issues with deep learning models is to set the hyperparameters. Using the considered strategy to the ensemble, the model has less dependency than the non-ensembled approach. Let n be the cardinality of models and m be the cardinality of all predictions made by each model. Formally, $\forall i = 1, ..., n$, $M = \{\mu_1, ..., \mu_n\}$, is the set of models; $P = \{\rho_1, ..., \rho_n\}$, is the set of all predictions made by M (i.e. μ_1 produces ρ_1), and $R_i = \max_{i,k}\{\rho_i, k \in P, k = 1, ..., m\}$, is the prediction that is maximum in P. Finally, $R = \bigcup_{k=1}^{m} R_i$, is the set of final predictions.

In detail, the predictions of all the stacked-CNN-BILSTMs are compared, and for each predicted tweet, the label is to be selected, which among the models is the maximum.

A. Experimental setup

The hyperparameters for the stacked-CNNs are illustrated in table:1, where for each CNN, there is a fixed filter (window size) and the same kernel and max-pooling. In table:1, the

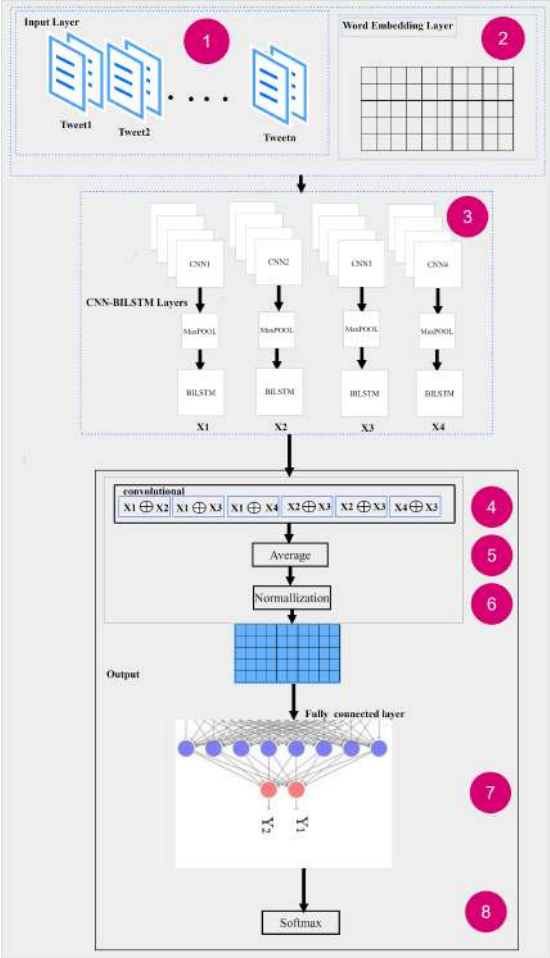


Fig. 1. Depiction of Multi-Stacked CNN-BILSTM

max-pool1 is used for each CNN while the max-pool2 is used for the merged CNNs. The BILSTMs have 128 nodes and a dropout of 0.20. The maximum number of words, the maximum length of the sequences, and the batch size are 10000, 400, and 256, respectively. In complex problems, to distinguish between positive and negative samples (i.e. non-rumor and rumor), linear functions do not work as well. Except for the last hidden layer that employs the activation function of softmax, other activations are relu. For our model, the pre-trained word embedding of 300-dimensional Glove is used, and the number of the hidden layers is 50. For optimization, the well-known optimizer of ADAM is used. For the implementation of machine learning and deep learning models, we have used Scikit-Learn, and Keras, respectively. The implementations were made in the google colab environment.

IV. Results

A. Evaluation Metrics

Standard criteria in eq:7, eq:8, eq:9, eq:10,eq:11 and eq:12 for the evaluation of the models are used. In the equations, TP indicates true positive, TN indicates true negative, FP indicates false positive, and FN indicates false negative. TP and FN point to the number of samples that are correctly and wrongly labeled as a non-rumor, respectively. Also, TN and FP point to the number of samples that are correctly and wrongly labeled as a rumor, respectively. Hence, for the rumor/non-rumor class, recall refers to the percentage of total predictions that correctly are classified, while precision is the total predictions made for the considered class. The F1 is the weighted harmonic mean of precision and recall. The accuracy is the ratio of the number of total correct predictions. For performing all experiments, we have employed 5-fold cross-validation with a validation set of 0.10.

$$Precision\ (rumor) = \frac{TP}{TP + FP} \tag{7}$$

$$Recall\ (rumor) = \frac{TP}{TP + FN} \tag{8}$$

$$Precision\ (non - rumor) = \frac{TN}{TP + FN} \tag{9}$$

$$Recall\ (non - rumor) = \frac{TN}{TP + FP} \tag{10}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{11}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{12}$$

B. Results

For ensembling, We have considered five variation models from multi-stacked-CNN-BILSTM, which are listed from 1 to 5. It should be pointed out that for the derived models, just hyperparameters that are added or removed is to be described. (1) The first model is exactly like what we have described so far. In a simple statement, the experimental setup noted, table:1, and word embedding 300-dimensional Glove. For this model, the reported accuracy is 86.80%. (2) This model is like the first model but instead of Glove employs Word2VEC, where the maximum number of words for Word2VEC is 10000. For this model, the maximum sequence length is 100, and the reported accuracy is 84.59%. (3) Similar to the first model but without any pre-trained word embedding, the reported accuracy is 86.07%. (4) Instead of capturing average and normalization, the outputs of four CNN-BILSTMS are merged, and the reported accuracy is 85.33%. (5) Instead of 50 hidden layers, 256 hidden layers, and the maximum number of words is 25000. The reported accuracy for this model is 85.97%. From the ensemble of models 1 to 5, the accuracy from 86.80 to 87.59 increased, which accuracy 86.80% is for multi-stacked-CNN-BILSTM and 87.59% is for the ensemble of multi-stacked-CNN-BILSTM. We have compared the proposed model (non-ensembled) and proposed ensembled model to some deep learning and machine learning models that were presented for rumor detection and text classification. Also, for investigating the robustness of the proposed models, we have implemented some popular machine learning algorithms for rumor detection. It is illustrated the detailed rumor detection results of each model in table:2. It can be seen that the proposed model (non-ensembled) and proposed ensembled model in terms of accuracy and f1 have superiority over the ten studied models. Except for our proposed models, the random forest has the best performance, where table:2 indicates that the accuracy of our ensembled model is 87.59 while the random forest is 86.34. In the proposed ensembled model, the correct predictions (recall) are more than rumor tweets. It is reasonable, as the number of non-rumor tweets is more than rumors. Also, table:2 indicates that f1 of the proposed model has superiority over the studied models. From the previously presented works, SVM-TS [23] for machine learning models, and CNN [22] for deep learning models have the best performance, where CNN outperforms to detect rumors than non-rumors while SVM-TS [23] outperforms to detect the non-rumors. Also, table:2 indicates that CSI [10] and BILSTM [9] have achieved very poor result and just can detect the rumors.

Also, our experimental results verify the considered hypotheses. In detail, for hypothesis 1, the multi-stacked-CNN-BILSTM outperformed the baseline models in terms of accuracy, and f1, which indicates the impact of considering local and global features extracted. Also, experimental results verify hypothesis 2, as the ensembled model outperforms the non-ensembled models in terms of accuracy and f1, where this outperforming indicates that the ensembled model doesn't depend on some specific hyperparameters.

TABLE II
COMPARISON OF RUMOR DETECTION RESULTS IN PHEME DATA SET

Model	Class	Accuracy	Precision	Recall	F1
CSI [10]	R N	33.98	34 0	1 0	51 0
RNN [8]	R N	49.12	34.20 66.23	53.60 46.78	41.60 54.82
LSTM [11]	R N	83.98	75.80 88.50	77.60 87.10	76.80 87.77
BILSTM [9]	R N	33.98	34 0	1 0	51 0
CNN [22]	R N	85.36	77 90.29	81.64 87.31	78.12 88.73
SVM-TS [23]	R N	85.48	79.46 88.46	77.28 89.71	78.35 89.08
SVM-BOW	R N	85.86	84.14 86.57	72.16 92.03	77.61 89.52
Random Forest-BOW	R N	86.34	85.13 86.85	72.51 93.46	78.29 90.03
Naive Bayes-BOW	R N	84.52	75.44 89.75	80.83 86.42	78.03 88.05
KNN-BOW	R N	78.81	84.08 77.65	46.65 95.37	59.90 85.59
Proposed Model	R N	**86.80**	81.97 89.12	78.40 91.12	**80.15** **90.11**
Proposed Ensembled Model	R N	**87.59**	84.25 89.13	78.09 92.48	**81.05** **90.77**

V. CONCLUSION

Social networks are widely available for the public, and they have many benefits, such as ease of communication, e-commerce, and keeping up to date with the latest global events. However, one of the negative factors in social networks is the existence of rumors that can have destructive financial and human effects. We pointed that how dangerous rumors like encouraging the consumption of chemical detergents to prevent coronavirus infection can be dangerous for people.

In this study, we have proposed an ensembled text-based multi-stacked-CNN-BILSTM for rumor detection. The employed ensemble is based on maximum vote such that the label that obtains the most votes among the models is to be selected as the predicted label. Usually, Rumors are written in a deceitful way that even experts cannot recognize. Hence, there is a necessity for models that can extract discriminative features. To capture the discriminative features of rumors and non-rumors, stacked-CNNs are utilized to extract local features, and BILSTMs are utilized to consider the dependency among the local features. It is utilized multiple Stacked-CNN-BILSTM to extract more discriminative features. It is discussed that ensemble deep learning became an active area of research and outperformed in some research fields. It is utilized a maximum-vote ensemble for multi-stacked-CNN-BILSTM, where the proposed model is ensembled with its variations. In our work, we have defined five variations of multi-stacked-CNN-BILSTM that increased accuracy from 86.80 to 87.59.

We discussed that hyperparameters of deep learning models impact the overall performance. Hence, it is noted that using an ensemble approach decreases the dependency of the deep learning model with its hyperparameters.

We have compared our proposed model over ten machine learning and deep learning models and found that the performance of our proposed model has superiority over the studied models. We achieved 87.59% accuracy, 81.05% f1 for detecting rumor tweets, and 90.77% f1 for detecting non-rumor tweets.

REFERENCES

[1] Rashid Behzadidoost, Mahdieh Hasheminezhad, Mohammad Farshi, Vali Derhami, and Farinaz Alamiyan-Harandi. A framework for text mining on twitter: a case study on joint comprehensive plan of action (jcpoa)-between 2015 and 2019. *Quality & Quantity*, pages 1–32, 2021.

[2] Adam Rapp, Lauren Skinner Beitelspacher, Dhruv Grewal, and Douglas E Hughes. Understanding social media effects across seller, retailer, and consumer interactions. *Journal of the Academy of Marketing Science*, 41(5):547–566, 2013.

[3] Alessandro Bondielli and Francesco Marcelloni. A survey on fake news and rumour detection techniques. *Information Sciences*, 497:38–55, 2019.

[4] Omer Sagi and Lior Rokach. Ensemble learning: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4):e1249, 2018.

[5] MA Ganaie, Minghui Hu, et al. Ensemble deep learning: A review. *arXiv preprint arXiv:2104.02395*, 2021.

[6] Yang Xing, Chen Lv, Huaji Wang, Dongpu Cao, and Efstathios Velenis. An ensemble deep learning approach for driver lane change intention inference. *Transportation Research Part C: Emerging Technologies*, 115:102615, 2020.

[7] Tao Zhou, Huiling Lu, Zaoli Yang, Shi Qiu, Bingqiang Huo, and Yali Dong. The ensemble deep learning model for novel covid-19 on ct images. *Applied Soft Computing*, 98:106885, 2021.

[8] Jing Ma, Wei Gao, Prasenjit Mitra, Sejeong Kwon, Bernard J Jansen, Kam-Fai Wong, and Meeyoung Cha. Detecting rumors from microblogs with recurrent neural networks. 2016.

[9] Akshi Kumar, Vaibhav Singh, Tuba Ali, Saurabh Pal, and Jeevanjot Singh. Empirical evaluation of shallow and deep classifiers for rumor detection. In *Advances in Computing and Intelligent Systems*, pages 239–252. Springer, 2020.

[10] Natali Ruchansky, Sungyong Seo, and Yan Liu. Csi: A hybrid deep model for fake news detection. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pages 797–806, 2017.

[11] Amin Saradar Torshizi and Adel Ghazikhani. Automatic twitter rumor detection based on lstm classifier. In *International Congress on High-Performance Computing and Big Data Analysis*, pages 291–300. Springer, 2019.

[12] Thanh Toan Nguyen, Thanh Tam Nguyen, Thanh Thi Nguyen, Bay Vo, Jun Jo, and Quoc Viet Hung Nguyen. Judo: Just-in-time rumour detection in streaming social platforms. *Information Sciences*, 570:70–93, 2021.

[13] Fan Xu, Victor S Sheng, and Mingwen Wang. Near real-time topic-driven rumor detection in source microblogs. *Knowledge-Based Systems*, 207:106391, 2020.

[14] Zhiyuan Wu, Dechang Pi, Junfu Chen, Meng Xie, and Jianjun Cao. Rumor detection based on propagation graph neural network with attention mechanism. *Expert systems with applications*, 158:113595, 2020.

[15] Pengfei Zhang, Hongyan Ran, Caiyan Jia, Xuanya Li, and Xueming Han. A lightweight propagation path aggregating network with neural topic model for rumor detection. *Neurocomputing*, 458:468–477, 2021.

[16] Oluwaseun Ajao, Deepayan Bhowmik, and Shahrzad Zargari. Fake news identification on twitter with hybrid cnn and rnn models. In *Proceedings of the 9th international conference on social media and society*, pages 226–230, 2018.

[17] Xiao Zheng, Wanzhong Chen, Yang You, Yun Jiang, Mingyang Li, and Tao Zhang. Ensemble deep learning for automated visual classification using eeg signals. *Pattern Recognition*, 102: 107147, 2020.

[18] Inwoong Lee, Doyoung Kim, Seoungyoon Kang, and Sanghoon Lee. Ensemble deep learning for skeleton-based action recognition using temporal sliding lstm networks. In *Proceedings of the IEEE international conference on computer vision*, pages 1012–1020, 2017.

[19] Sai Ma and Fulei Chu. Ensemble deep learning-based fault diagnosis of rotor bearing systems. *Computers in industry*, 105: 143–152, 2019.

[20] Ana Fidalgo, Ibéria Medeiros, Paulo Antunes, and Nuno Neves. Towards a deep learning model for vulnerability detection on web application variants. In *2020 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, pages 465–476. IEEE, 2020.

[21] Ronald J Williams and David Zipser. Gradient-based learning algorithms for recurrent. *Backpropagation: Theory, architectures, and applications*, 433:17, 1995.

[22] Yoon Kim. Convolutional neural networks for sentence classification. In *EMNLP*, 2014.

[23] Jing Ma, Wei Gao, Zhongyu Wei, Yueming Lu, and Kam-Fai Wong. Detect rumors using time series of social context information on microblogging websites. In *Proceedings of the 24th ACM international on conference on information and knowledge management*, pages 1751–1754, 2015.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Application of fuzzy generalized power series for the fuzzy linear inhomogeneous differential equation

1st S. Zabihi, 2nd R. Ezzati

1st Department of Mathematics, Central Tehran Branch, Islamic Azad University, Tehran, Iran, samanehzabihi95@yahoo.com
2nd Department of Mathematics, Karaj Branch, Islamic Azad University, Karaj, Iran, ezati@kiau.ac.ir

***Abstract*— In this paper, fuzzy linear Inhomogeneous differential equation under generalized Hukuhara differentiability with real coefficients is solved. For this purpose, we used of fuzzy generalized power series to solved this problem. Since in the discussion of fuzzy generalized power series, the concept of the fuzzy absolute convergence is one of the most important and fundamental concepts. In addition, an ordinary points, singular points and irregular singular points of the fuzzy second-order differential equation with a real coefficients are introduced. To complete the discussion, it has been shown that uniqueness of the solution of the fuzzy linear inhomogeneous second-order differential equations with real coefficients in the form of a fuzzy generalized power series by attention to the type of generalized differentiability is shown.**

***Index Terms*—Fuzzy Analytic Function, Fuzzy Generalized Power Series, Fuzzy Absolute Convergence, Fuzzy Linear Inhomogeneous Differential Equation.**

I. Introduction

In recent decades, there has been a great development of fuzzy concepts and the possibility of modeling phenomena taking into account the vagueness and uncertainties existing, therefore the proposed model is more consistent with existing reality and can express issues with a more comprehensive view.

The first definition of the fuzzy differential equation and the fuzzy initial value problem was given by Nieto and Rodriguez-Lopez , Seikkala and then extended and used in many research articles [1–3, 6, 7, 12, 16, 18, 20–22]. Due to the development of the theory of fuzzy differential equations, solving this kind problems by using numerical methods have been increased [8, 9]. One of the applicable method for solving fuzzy differential equations with variable coefficients is the fuzzy series method when we want to solve the equations at an ordinary point. Kadak in [14] introduced the power series of fuzzy numbers and in [15] using the H-differentiation and radius of convergence, the relationship between term by term H-differentiation and uniform convergence of fuzzy valued function series have been shown. Sabzi in [19] introduced the fuzzy generalized power series method under generalized Hukuhara differentiability, fuzzy comparison test, Ratio Test and solved fuzzy Legendre differential equation by fuzzy generalized power series method.

In the present work, we apply the fuzzy generalized power series method to solve the fuzzy Linear Inhomogeneous Differential equation under generalized Hukuhara differentiability with real coefficients. The paper is organized as follows. At first, in Section II we express briefly basic concepts and theorems with introducing notation also, we prove some new theorems and lemmas to be used in the main part of the paper. Section III investigates the basic concepts and theorems of fuzzy generalized power series under $gH-$differentiability, fuzzy analytic functions, ordinary and singular points, fuzzy series and uniqueness and convergence of fuzzy differential equation by fuzzy generalized power series. Section IV some concept and important theorem such as fuzzy absolute convergence are proved. Section V We show that the fuzzy inhomogeneous differential equation equation with real coefficients has a unique solution in the form of the fuzzy generalized power series by attention to the type of $gH-$differentiability, conclusions are drawn in Section VI.

II. Fuzzy Basic Operation

We denote by $\mathbb{R}_{\mathcal{F}}$, the set of fuzzy numbers, that is, normal, fuzzy convex, upper semi-continuous and compactly supported fuzzy sets which defined over the real line. Let a is a fuzzy number. For $0 < \alpha \leq 1$, $[u]^{\alpha} = [u^{-}, u^{+}]$ is the α-cut representation for this fuzzy number.

Definition II.1. *(See [17]) Let $u, v, c, e \in \mathbb{R}_{\mathcal{F}}$ and $\lambda \in \mathbb{R}$. The metric D on $\mathbb{R}_{\mathcal{F}}$ given by $D : \mathbb{R}_{\mathcal{F}} \times \mathbb{R}_{\mathcal{F}} \longrightarrow \mathbb{R}^{+} \cup \{0\}$*

$$D(u,v) = \sup_{0\leq\alpha\leq 1} d_H\Big([u]^{\alpha}, [v]^{\alpha}\Big) = \sup_{0\leq\alpha\leq 1} \max\Big\{|u^{-}-v^{-}|, |u^{+}-v^{+}|\Big\}.$$

where $[u]^{\alpha} = [u^{-}, u^{+}]$, $[v]^{\alpha} = [v^{-}, v^{+}]$ are α-cut's of u and v, and d_H is the Hausdorff metric. D is a complete metric in $\mathbb{R}_{\mathcal{F}}$ and has the following properties

1. $D(u\oplus c, v\oplus c) = D(u,v)$,
2. $D(\lambda \odot u, \lambda \odot v) = |\lambda| D(u,v)$,
3. $D(u\oplus v, c\oplus e) \leq D(u,c) + D(v,e)$.

Definition II.2. *(See [13]) Let $u, v \in \mathbb{R}_{\mathcal{F}}$ and $k \in \mathbb{R}$, $([u]^{\alpha} = [u^-, u^+],\ [v]^{\alpha} = [v^-, v^+])$, then for all $\alpha \in [0,1]$*

1). $u \succ 0$ *if and only if* $0 < u^- \leq u^+$,
2). $u \prec 0$ *if and only if* $u^- \leq u^+ < 0$,
3). $u \succeq v$ *if and only if* $(v^- \leq u^-$ *and* $v^+ \leq u^+)$,
4). $u \prec v$ *if and only if* $(v^- < u^-$ *and* $v^+ < u^+)$,
5). $k \preceq v$ *if and only if* $(k \leq v^-$ *and* $k \leq v^+)$,
6). $0 \in u$ *if and only if* $u^- \leq 0,\ u^+ \geq 0$.

The set of all positive fuzzy numbers is denoted by $\mathbb{R}^+_{\mathcal{F}}$.

Definition II.3. *(See [11]) Consider $u, v \in \mathbb{R}_{\mathcal{F}}$. If there exists $w \in \mathbb{R}_{\mathcal{F}}$ so that $u = v + w$, then w is called the Hukuhara difference of u and v, and it is defined by $u \ominus v$. Let $u \ominus v$ exists, it is unique and its $\alpha-$cut's are $[u \ominus v]^{\alpha} = [u^- - v^-, u^+ - v^+]$.*

Definition II.4. *(See [21]) The generalized Hukuhara difference of two fuzzy number $u, v \in \mathbb{R}_{\mathcal{F}}$ is denoted as follows*

$$u \ominus_{gH} v = w \Longleftrightarrow \begin{cases} (i)\ \ u = v + w, \\ or \\ (ii)\ \ v = u + (-1)w. \end{cases}$$

The fuzzy $gH-$difference $\ominus_{gH}$ is well specific if the α-cuts $[w]^{\alpha}$ are such that $w \in \mathbb{R}_{\mathcal{F}}$ (*i.e.*, w^- is nondecreasing, w^+ is non increasing and $w^- \leq w^+$).

Definition II.5. *(See [21]) Consider $u, v \in \mathbb{R}_{\mathcal{F}}$ be two fuzzy numbers. Then*

(i) *if the $gH-$difference exists, it is unique,*
(ii) $u \ominus_{gH} v = u \ominus v$ *or* $u \ominus_{gH} v = -(v \ominus u)$ *whenever the explanations on the right exist, in particular,* $u \ominus_{gH} u = u \ominus u = 0$,
(iii) *if $u \ominus_{gH} v$ exists in the sense (i), then $v \ominus_{gH} u$ exists, in the sense (ii) and vice versa,*
(iv) $(u + v) \ominus_{gH} v = u$,
(v) $0 \ominus_{gH} (u \ominus_{gH} v) = v \ominus_{gH} u$,
(vi) $u \ominus_{gH} v = v \ominus_{gH} u = w$ *if and only if* $w = -w$. *Furthermore,* $w = 0$ *if and only if* $u = v$.

Definition II.6. *(See [10]) A fuzzy-valued function $f : [a,b] \to \mathbb{R}_{\mathcal{F}}$ is said to be continuous at $x_0 \in [a,\ b]$ if for each $\epsilon > 0$ there is $\delta > 0$ such that $D\Big(f(x), f(x_0)\Big) < \epsilon$, whenever $x \in [a,\ b]$ and $|x - x_0| < \delta$. We say that f is fuzzy continuous on $[a,b]$ if f is continuous at each $x_0 \in [a,\ b]$.*

Lemma II.1. *(See [19]) If $u \in \mathbb{R}_{\mathcal{F}}$ and μ_1, μ_2 be two real numbers. Then*

(i) *If $\mu_1 - \mu_2 \geq 0$, then*

$$\ominus(-1)\mu_1 u \oplus (-1)\mu_2 = \ominus(-1)(\mu_1 - \mu_2)u$$

(ii) *If $\mu_1 - \mu_2 \leq 0$, then*

$$\ominus(-1)\mu_1 u \oplus (-1)\mu_2 = (\mu_1 - \mu_2)u$$

Definition II.7. *(See [21]) The generalized division $\div_g$ of two fuzzy numbers u and v we define as follows*

$$u \div_g v = w \Longleftrightarrow \begin{cases} (i)\ \ u = vw, \\ or \\ (ii)\ \ v = uw^{-1}. \end{cases}$$

Definition II.8. *Let $f_n(x) = a_n^+(0)x^n$ such that for all $n \in N \cup \{0\}$, we have $a_n^+(0) \in \mathbb{R}$ and $x^n \in \mathbb{R}$. If a function $f_n : (c, +\infty) \to \mathbb{R}^+$ is defined on $(c, +\infty)$, such that $\lim_{x\to\infty} f_n(x) = h$, then for all $\delta > 0$, $\exists N > 0$ such that $D\Big(f(x), h\Big) < \delta$ (for all $x > N$).*

Proposition II.2. *If a_n and a_{n+1} are positive fuzzy number then*

$$a_{n+1} \div_g a_n = a_{n+1} \div a_n.$$

Proof. *According to the assumptions of this Proposition $a_n \in \mathbb{R}^+_{\mathcal{F}}$ and $a_{n+1} \in \mathbb{R}^+_{\mathcal{F}}$, so according to Definition II.2 (1), we have $0 < a_n^- \leq a_n^+$ and $0 < a_{n+1}^- \leq a_{n+1}^+$, then by using Definition II.7 (ii), the requited result is obtained.* □

Definition II.9. *(See [5]) The generalized Hukuhara derivative ($gH-$derivative) of a fuzzy-valued function $f : (a,b) \longrightarrow \mathbb{R}_{\mathcal{F}}$ at x_0 is defined as*

$$f'_{gH}(x_0) = \lim_{h\to 0} \frac{f(x_0 + h)\ \ominus_{gH}\ f(x_0)}{h},$$

if $f'_{gH}(x_0) \in \mathbb{R}_{\mathcal{F}}$, we say that f is generalized Hukuhara differentiable ($gH-$differentiable) at x_0.
Also we say that f is $[(i) - gH]-$differentiable at x_0 if

$$[f'_{gH}(x_0)]^{\alpha} = [f'^{-}(x_0), f'^{+}(x_0)], \qquad 0 \leq \alpha \leq 1, \quad (1)$$

and that f is $[(ii) - gH]-$differentiable at x_0 if

$$[f'_{gH}(x_0)]^{\alpha} = [f'^{+}(x_0), f'^{-}(x_0)], \qquad 0 \leq \alpha \leq 1. \quad (2)$$

Definition II.10. *(See [5]) If $f : (a,b) \longrightarrow \mathbb{R}_{\mathcal{F}}$ and $f(x)$ is $gH-$differentiable of the order j, $j = 1, 2, ..., n-1$, at x_0, provided that $gH-$differentiable type has no change, then there exist $f^{(n)}_{gH}(x_0) \in \mathbb{R}_{\mathcal{F}}$ such that*

$$f^{(n)}_{gH}(x_0) = \lim_{h\to 0} \frac{f^{(n-1)}_{gH}(x_0 + h)\ \ominus_{gH}\ f^{(n-1)}_{gH}(x_0)}{h}.$$

Definition II.11. *(See [14]) If $x \in \mathbb{R}$ be any element and $x_0 \in \mathbb{R}$ the fixed element, then the power series with fuzzy coefficients a_n is in the following equation*

$$\sum_{n=0}^{\infty} a_n \odot (x - x_0)^n = a_0 \oplus a_1 \odot (x - x_0)^1 \oplus \cdots \oplus a_n \odot (x - x_0)^n \oplus \cdots .$$

III. Fuzzy Generalized Power Series and Fuzzy Analytic Functions

In this section, we are introduce basic concepts of fuzzy generalized power series under $gH-$differentiability, fuzzy analytic functions, ordinary and singular points, fuzzy sreies and uniqueness and convergence of fuzzy differential equation by fuzzy generalized power series. These concepts are basics for Sections IV and V.

Definition III.1. *(Fuzzy analytic functions) The fuzzy valued function $f(x)$ is said to be fuzzy analytic at $x = x_0$, if $f(x)$ in an open interval around x_0, has a fuzzy Taylor expansion, with a positive radius of convergence ($\rho > 0$).*

Definition III.2. *(See [19])(Fuzzy generalized power series) Suppose that $x, x_0 \in \mathbb{R}$ and $a_n \in \mathbb{R}_{\mathcal{F}}$ such that $a_n :=$*

$\frac{f_{gH}^{(n)}(x_0)}{n!}$, *we introduce* $f(x)$ *the fuzzy generalized power series as shown by*

$$f(x) = \sum_{n=0}^{\infty} a_n \odot (x - x_0)^n.$$

Theorem III.1. *(See [19]) Let* $f(x) = \sum_{n=0}^{\infty} a_n \odot (x-x_0)^n$ *be a fuzzy generalized power series then we have*

Type(i) $f(x) = \sum_{n=0}^{\infty} a_n \odot (x - x_0)^n$,
Type(ii) $f(x) = a_0 \ominus (-1) \sum_{n=1}^{\infty} a_n \odot (x - x_0)^n$,
Type(iii) $f(x) = \sum_{n=2k} a_n \odot (x - x_0)^n \ominus (-1) \sum_{n=2k+1} a_n \odot (x - x_0)^n$.

Theorem III.2. *(See [19]) Let* $f(x) = \sum_{n=0}^{\infty} a_n \odot (x - x_0)^n$ *(fuzzy power series) with* $\rho > 0$*, then* $f(x)$ *is* $gH-$*differentiable in* $|x - x_0| < \rho$*, and we have*

$$f'_{gH}(x) = \sum_{n=0}^{\infty} n \odot a_n \odot (x - x_0)^{n-1}. \tag{3}$$

with the same radius of convergence.

Definition III.3. *(See [19]) (See [14]) The radius of convergence* ρ *is defined by*

$$\rho := \lim_{n\to\infty} \frac{a_n^+(0)}{a_{n+1}^+(0)},$$

which is also given provided that the limit on the right hand side exists, where $0 \le \rho \le \infty$.

Theorem III.3. *(See [19]) If* $f(x)$ *and* $g(x)$ *are fuzzy generalized power series with* $\rho > 0$ *and* $q(x)$ *is real power series with* $\rho > 0$,

(1) *then the multiplication between a fuzzy generalized power series and real power series is obtained*
   (i) $f(x) \odot q(x) = \sum_{n=0}^{\infty} \left[\sum_{k=0}^{n} (a_k \odot c_{n-k})\right] \odot (x - x_0)^n$,
   (ii) $f(x) \odot q(x) = \sum_{n=0}^{\infty} \left[(a_0 \odot c_n) \ominus (-1) \sum_{k=1}^{n} (a_k \odot c_{n-k})\right] \odot (x - x_0)^n$,
   (iii) $f(x) \odot q(x) = \sum_{n=0}^{\infty} \left[(a_0 \odot c_n) \oplus \sum_{k=1}^{n} (\ominus)^k (-1)^k (a_k \odot c_{n-k})\right] \odot (x - x_0)^n$.

(2) *then the addition of two fuzzy generalized power series defined by*
   (i) $f(x) \oplus g(x) = \sum_{n=0}^{\infty} (a_n \oplus b_n) \odot (x - x_0)^n$,
   (ii) $f(x) \oplus g(x) = (a_0 \oplus b_0) \ominus (-1) \sum_{n=1}^{\infty} (a_n \oplus b_n) \odot (x - x_0)^n$,
   (iii) $f(x) \oplus g(x) = \sum_{i=2k} (a_i \oplus b_i) \odot (x - x_0)^i \oplus \sum_{i=2k+1} \ominus(-1)(a_i \oplus b_i) \odot (x - x_0)^i$.

(3) *then the subtraction of two fuzzy generalized power series defined by*
   (i) $f(x) \ominus g(x) = \sum_{n=0}^{\infty} (a_n \ominus b_n) \odot (x - x_0)^n$,
   (ii) $f(x) \ominus g(x) = (a_0 \ominus b_0) \ominus (-1) \sum_{n=1}^{\infty} (a_n \ominus b_n) \odot (x - x_0)^n$,
   (iii) $f(x) \ominus g(x) = \sum_{i=2k} (a_i \ominus b_i) \odot (x - x_0)^i \oplus \sum_{i=2k+1} \ominus(-1)(a_i \ominus b_i) \odot (x - x_0)^i$.

Theorem III.4. *(See [19]) Let* $f(x)$ *be a fuzzy generalized power series with* $\rho > 0$*, then* $f'_{gH}(x)$ *is* $gH-$*differentiable in* $|x - x_0| < \rho$*, with the same radius of convergence and we have*

1. *If* $f(x)$ *be a fuzzy generalized power series of type (i), then*
   I. $f'_{i-gH}(x) = \sum_{n=1}^{\infty} a_n \odot n \odot (x - x_0)^{n-1}$,
   II. $f'_{ii-gH}(x) = \ominus(-1) \sum_{n=1}^{\infty} a_n \odot n \odot (x - x_0)^{n-1}$,
2. *If* $f(x)$ *be a fuzzy generalized power series of type (ii), then*
   I. $f'_{i-gH}(x) = \ominus(-1) \sum_{n=1}^{\infty} a_n \odot n \odot (x - x_0)^{n-1}$,
   II. $f'_{ii-gH}(x) = \sum_{n=1}^{\infty} a_n \odot n \odot (x - x_0)^{n-1}$.
3. *If* $f(x)$ *be a fuzzy generalized power series of type (iii), then*
   I. $f'_{i-gH}(x) = \ominus(-1) \sum_{n=2k} (n+1) \odot a_{n+1} \odot (x - x_0)^n \oplus \sum_{n=2k+1} (n+1) \odot a_{n+1} \odot (x - x_0)^n$,
   II. $f'_{ii-gH}(x) = \sum_{n=2k} (n+1) \odot a_{n+1} \odot (x-x_0)^{n-1} \ominus (-1) \sum_{n=2k+1} (n+1) \odot a_{n+1} \odot (x - x_0)^{n-1}$.

With the same radius of convergence.

Theorem III.5. *(See [19]) Let fuzzy generalized power series* $f(x)$ *with* $\rho > 0$*, then* $f(x)$ *is* $gH-$*differentiable in* $|x - x_0| < \rho$*, and* $f''_{gH}(x)$ *defined by one of the following states is expressed*

with the same radius of convergence and we have

1. *If* $f(x)$ *be a fuzzy generalized power series of type (i), then*
   I. $f''_{i-gH}(x) = \sum_{n=1}^{\infty} n \odot (n-1) \odot a_n \odot (x - x_0)^{n-2}$,
   II. $f''_{ii-gH}(x) = \ominus(-1) \sum_{n=1}^{\infty} n \odot (n-1) \odot a_n \odot (x - x_0)^{n-2}$.
2. *If* $f(x)$ *be a fuzzy generalized power series of type (ii), then*
   I. $f''_{i-gH}(x) = \ominus(-1) \sum_{n=1}^{\infty} n \odot (n-1) \odot a_n \odot (x - x_0)^{n-2}$,
   II. $f''_{ii-gH}(x) = \sum_{n=1}^{\infty} n \odot (n-1) \odot a_n \odot (x-x_0)^{n-2}$.
3. *If* $f(x)$ *be a fuzzy generalized power series of type (iii), then*
   I. $f''_{i-gH}(x) = \sum_{n=2k} (n+1)(n+2) \odot a_{n+2} \odot (x - x_0)^n \ominus (-1) \sum_{n=2k+1} (n+1)(n+2) \odot a_{n+2} \odot (x - x_0)^n$,
   II. $f''_{ii-gH}(x) = \ominus(-1) \sum_{n=2k} (n+1)(n+2) \odot a_{n+2} \odot (x - x_0)^n \oplus \sum_{n=2k+1} (n+1)(n+2) \odot a_{n+2} \odot (x - x_0)^n$.

With the same radius of convergence.

A. Ordinary and Singular Points

Definition III.4. *Consider* $f : [a, b] \to \mathbb{R}_{\mathcal{F}}^+$ *and* $x, x_0 \in \mathbb{R}$. *The fuzzy valued function* $f(x)$ *is said to be positive fuzzy analytic at* $x = x_0$*, if* $f(x)$ *in an open interval around* x_0*, has a fuzzy Taylor expansion with a fuzzy positive radius of convergence* $\rho_m \succ 0$.

Definition III.5. *(Ordinary and singular points of the FSODE) Let* $f : [a, b] \to \mathbb{R}_{\mathcal{F}}^+$ *and* $R, P, Q : [a, b] \to \mathbb{R}$. *Consider the*

following fuzzy second-order differential equation with a real coefficients (FSODE)

$$\begin{cases} R(x) \odot f''_{gH}(x) \oplus P(x) \odot f'_{gH}(x) \oplus Q(x) \odot f(x) = 0, \\ \\ f(x_0) = f_0 \in \mathbb{R}_{\mathcal{F}}, \qquad f'_{gH}(x_0) = f_1 \in \mathbb{R}_{\mathcal{F}}. \end{cases} \tag{4}$$

If $R(x), P(x), Q(x)$ are real analytic at $x = x_0$ and $R(x_0) \neq 0$ then x_0 is called an ordinary point of the Eq. (4) *(FSODE). If x_0 is not an ordinary point then it is called a singular point of the Eq.* (4) *(FSODE).*

Remark III.6. *In Eq.* (4) *we assume that $R(x)$, $P(x)$, $Q(x)$ are real valued functions and $R(x_0) \neq 0$ then, we get $p(x) = P(x) \div R(x)$ and $q(x) = Q(x) \div R(x)$*

Remark III.7. *Let $p, q : [a, b] \to \mathbb{R}$ and $x, x_0 \in \mathbb{R}$. If $\lim_{x \to x_0} p(x)$ and $\lim_{x \to x_0} q(x)$ exist at point x_0 it is called an ordinary point of FSODE, but if at least one of two limits, doesn't then the point $x = x_0$ is a singular point of FSODE.*

Definition III.6. *(Singular and irregular singular points of the FSODE) Let $p, q : [a, b] \to \mathbb{R}$ and $x, x_0 \in \mathbb{R}$. If x_0 is a singular point then we get two cases*

(a). *If $\lim_{x \to x_0} p(x)(x - x_0)$ and $\lim_{x \to x_0} q(x)(x - x_0)^2$ exist at point x_0 then x_0 is regular singular of FSODE.*

(b). *If at least one of two limits $p(x)(x - x_0)$ and $q(x)(x - x_0)^2$ doesn't at point x_0 then x_0 is irregular singular of FSODE.*

Definition III.7. *(See [19])(Ordinary and singular points of the FSSODE) Suppose that $f(x)$ is fuzzy value functions and $Q(x), P(x)$ and $R(x)$ are real value functions such that $R(x)$ is coefficient of $f''(x)$. Consider the following fuzzy second-order differential equation in standard form (FSSODE)*

$$\begin{cases} f''_{gH}(x) \oplus P(x) \odot f'_{gH}(x) \oplus Q(x) \odot f(x) = 0, \\ \\ f(x_0) = f_0 \in \mathbb{R}_{\mathcal{F}}, \qquad f'_{gH}(x_0) = f_1 \in \mathbb{R}_{\mathcal{F}}. \end{cases} \tag{5}$$

If $R(x), P(x), Q(x)$ are real analytic function at $x = x_0$ and also $R(x_0) \neq 0$ then x_0 is called an ordinary point of the Eq. (5)*. If x_0 is not an ordinary point then it is called a singular point of the Eq.* (5)*.*

Remark III.8. *(See [19]) Assume that. If both $\lim_{x \to x_0} p(x)$ and $\lim_{x \to x_0} q(x)$ exist at point x_0 it is called an ordinary point, but if either $\lim_{x \to x_0} p(x)$ or $\lim_{x \to x_0} q(x)$ doesn't exist, then the point $x = x_0$ is a singular point.*

Remark III.9. *(See [19]) Let $x, x_0 \in \mathbb{R}$ and $R(x), P(x), Q(x)$ are real analytic function at $x = x_0$. If x_0 is a singular point then we have two cases*

- *If $p(x)(x - x_0)$ and $q(x)(x - x_0)^2$ are analytic, then x_0 is called regular singular;*
- *Otherwise x_0 is called irregular singular.*

B. Fuzzy Sreies

Definition III.8. *(See [19]) u is called fuzzy sequence if For all $n \in N$, we have $u = \{u_n\}$ with $u : Z^+ \to \mathbb{R}_{\mathcal{F}}$.*

Definition III.9. *(See [19]) u is called fuzzy positive sequence if For all $n \in N$, we have $u = \{u_n\}$ with $u : Z^+ \to \mathbb{R}^+_{\mathcal{F}}$.*

Definition III.10. *(See [19]) An infinite fuzzy series $\sum_{n=1}^{\infty} u_n$ with a positive terms is convergent if and only if its fuzzy sequence of partial sums has an fuzzy upper bound.*

Remark III.10. *If $\sum_{n=0}^{\infty} a_n \odot (x - x_0)^n$ is a fuzzy generalized power series then for all x*

i. *This fuzzy generalized power series is converges if $|x - x_0| \prec \rho_m$,*
ii. *This fuzzy generalized power series is diverges if $|x - x_0| \succ \rho_m$.*

Definition III.11. *(See [19]) Assume that the fuzzy generalized power series $\sum_{k=0}^{\infty} a_n \odot x^n$ having a radius of convergence of $\rho > 0$. For each arbitrary number x_i such that $|x_i| < \rho$, we have $\lim_{n \to \infty} a_n \odot x_i^n = 0$.*

Definition III.12. *Let $f_n(x) = a_n^+(0)x^n$ such that for all $n \in N \cup \{0\}$, we have $a_n^+(0) \in \mathbb{R}$ and $x^n \in \mathbb{R}$. If a function $f_n : (c, +\infty) \to \mathbb{R}^+$ is defined on $(c, +\infty)$, such that $\lim_{x \to \infty} f_n(x) = h$, then for all $\delta > 0$, $\exists N > 0$ such that $D(f(x), h) < \delta$ (for all $x > N$).*

Proposition III.11. *If $f(x) : [a, b] \to \mathbb{R}^+_{\mathcal{F}}$, then $D(f(x), 0) = f(x)$ and $D(\ominus f(x), 0) = f(x)$.*

Proof. If $f(x) : [a, b] \to \mathbb{R}^+_{\mathcal{F}}$, then according to Definition II.2 (1), we have $0 \prec f^-(x) \preceq f^+(x)$, thus $D([f^-(x), f^+(x)], 0)$, it means $D(f^-(x), 0) = f^-(x)$ and $D(f^+(x), 0) = f^+(x)$, Then $D([f^-(x), f^+(x)], 0) = [f^-(x), f^+(x)]$, so $D(f(x), 0) = f(x)$.
We know that $\ominus$ is a binary operator then by using to Definition II.3 and using $\alpha-$cut's, we have $\ominus f(x)$ it means $\ominus[f^-(x), f^+(x)] = 0 \ominus [f^-(x), f^+(x)] = [-f^-(x), -f^+(x)]$, therefore $\ominus f(x) = [-f^-(x), -f^+(x)]$. Thus $D(\ominus f(x), 0) = D(0 \ominus [f^-(x), f^+(x)], 0) = D([-f^-(x), -f^+(x)], 0)$, it means $D(-f^-(x), 0) = f^-(x)$ and $D(-f^+(x), 0) = f^+(x)$, therefore $D(\ominus f(x), 0) = f(x)$ □

Theorem III.12. *(See [19])(Fuzzy Comparison Test) If $\sum_{n=1}^{\infty} u_n$ and $\sum_{n=1}^{\infty} v_n$ are two fuzzy series with a fuzzy positive terms such that for all n, $n \in N$, $u_n \succeq v_n$, then*

(i) *If a fuzzy series $\sum_{n=1}^{\infty} u_n$ is convergent, then a fuzzy series $\sum_{n=1}^{\infty} v_n$ is also convergent.*
(ii) *If a fuzzy series $\sum_{n=1}^{\infty} u_n$ is divergent, then a fuzzy series $\sum_{n=1}^{\infty} v_n$ is also divergent.*

Theorem III.13. *(See [19])(Ratio Test) Let $\sum_{k=0}^{\infty} u_n$ is the fuzzy series with positive terms, (For all $n \in N \cup \{0\}$, $u_n := a_n x^n$, $a_n \in \mathbb{R}^+_{\mathcal{F}}, x^n \in \mathbb{R}^+$), such that*

$$L_{RT} = \lim_{n \to \infty} D(\frac{u^+_{n+1}(0)}{u^+_n(0)}, 0) = \lim_{n \to \infty} D(\frac{a^+_{n+1}(0)x^{n+1}}{a^+_n(0)x^n}, 0). \tag{6}$$

(i) *If $L_{RT} < 1$, then the fuzzy series is convergent.*
(ii) *If $L_{RT} > 1$, then the fuzzy series is divergent.*

C. *Uniqueness and Convergence*

Theorem III.14. *(See [19])(Uniqueness analytic solution) If $T(x)$, $W(x)$ are real analytic functions at x_0 on an interval $|x-x_0|<\rho,\ \rho>0$ and let $f(x)$ is a fuzzy valued function and x_0 is an ordinary point of following fuzzy linear differential equation*

$$\begin{cases} f''_{gH}(x)\oplus T(x)\odot f'_{gH}(x)\oplus W(x)\odot f(x)=0,\\ f(x_0)=a_0\in\mathbb{R}_{\mathcal{F}},\\ f'_{gH}(x_0)=a_1\in\mathbb{R}_{\mathcal{F}}. \end{cases} \tag{7}$$

Then there exists a unique fuzzy solution

$$f(x)=\sum_{n=0}^{\infty}a_n\odot x^n,$$

under $[(i)-gH]-$differentiable, which is fuzzy analytic on the same interval.

Theorem III.15. *(See [19]) Let $T(x)$, $W(x)$ are real analytic functions at x_0 on $|x-x_0|<\rho,\ \rho>0$, $f(x)$ is fuzzy valued function and x_0 is an ordinary point of* (7)*, then under $[(i)-gH]-$differentiable there exists a unique following fuzzy solution*

$$f(x)=a_0\ominus(-1)\sum_{n=0}^{\infty}a_n\odot x^n.$$

which is fuzzy analytic on the same interval.

Theorem III.16. *(See [19]) Suppose that $T(x)$, $W(x)$ are real analytic functions at x_0 on $|x-x_0|<\rho,\ \rho>0$ and and $f(x)$ is a fuzzy valued function and also let x_0 is an ordinary point of the fuzzy linear differential equation* (7)*. then under $[(i)-gH]-$differentiability there exists a unique fuzzy solution*

$$f(x)=\sum_{n=2k}a_n\odot x^n\ominus(-1)\sum_{n=2k+1}a_n\odot x^n.$$

which is fuzzy analytic on the same interval.

IV. New Properties of Fuzzy Generalized Power Series

This section, some important definitions and theorems such as fuzzy absolute convergence and the same of radius of covergence of $f(x)$ and $f'_{gH}(x)$ will be proved.

Definition IV.1. *(Fuzzy Absolute Convergence)*
The fuzzy generalized power series $\sum_{n=0}^{\infty}a_n\odot(x-x_0)^n$ is said to absolutely converges if the fuzzy generalized power series $\sum_{n=0}^{\infty}|a_n\odot(x-x_0)^n|$ converges.

Definition IV.2. *The nth partial sums of the fuzzy generalized power series $\sum_{n=1}^{\infty}a_n\odot x^n$, is $\{s_n\}$. If $\sum_{n=1}^{\infty}a_n\odot x^n$ is convergent, then $\lim_{n\to\infty}s_n$ exist.*

Lemma IV.1. *Suppose that $z\neq 0$ is any real constant and $f(x)=\sum_{n=1}^{\infty}a_n\odot x^n$ is fuzzy generalized power series, then*
(a) *If the fuzzy generalized power series $\sum_{n=1}^{\infty}a_n\odot x^n$ is convergent and it is sum is S, then the fuzzy generalized power series $\sum_{n=1}^{\infty}za_n\odot x^n$ is convergent and it is sum is Sz.*
(b) *If the fuzzy generalized power series $\sum_{n=1}^{\infty}a_n\odot x^n$ is divergent, then the fuzzy generalized power series $\sum_{n=1}^{\infty}za_n\odot x^n$ is divergent.*

Proof. Suppose that the sum n of the first is the fuzzy generalized power series $\sum_{n=1}^{\infty}a_n\odot x^n$ be s_n. So $s_n=a_1\odot x^1\oplus a_2\odot x^2\oplus\ldots\oplus a_n\odot x^n$ and also nth partial sum of the fuzzy generalized power series $\sum_{n=1}^{\infty}za_n\odot x^n$ is $za_1\odot x^1\oplus za_2\odot x^2\oplus\ldots\oplus za_n\odot x^n=z(a_1\odot x^1\oplus a_2\odot x^2\oplus\ldots\oplus a_n\odot x^n)=zs_n$. Hence if the fuzzy generalized power series is convergent, then according to Definition IV.2 $\lim_{n\to\infty}s_n$ exist and is S, so

$$\lim_{n\to\infty}z\odot s_n=z\lim_{n\to\infty}s_n=zS$$

Therefor the fuzzy generalized power series $\sum_{n=1}^{\infty}z\odot a_n\odot x^n$ is convergent. □

Theorem IV.2. *Assume that there is Hukuhara difference. If the fuzzy generalized power series $f(x)=\sum_{n=0}^{\infty}a_n\odot(x-x_0)^n$ has radius of convergence $\rho_m>0$ then, $f(x)$ is $gH-$differentiable on $|x-x_0|<\rho_m$ and $f'_{gH}(x)=\sum_{n=0}^{\infty}na_n\odot(x-x_0)^{n-1}$, also has $\rho_m>0$ as its radius of convergence.*

Proof. Suppose without loss of generality that $x_0=0$. Assume x,x_1 be a numbers such that $|x|<|x_1|<\rho_m$, then according Remark III.10 (i), the fuzzy generalized power series $\sum_{n=0}^{\infty}a_n\odot x_1^n$ is convergent, therefor by Definition III.11, we have $\lim_{x\to\infty}a_n\odot x_1^n=0$. By using Definition III.12 we have $D(a_n\odot x_1^n,0)<\delta$, by considering $\delta=1$, we obtain $D(a_n\odot x_1^n,0)<1$.

$$\begin{aligned} D(na_n\odot x^{n-1},0)&=D(na_n\odot x^{n-1}\frac{x_1^n}{x_1^n},0.\frac{x_1^n}{x_1^n})\\ &=D(na_n\odot\frac{x^{n-1}}{x_1^{n-1}}\frac{x_1^n}{x_1},0.\frac{x_1^n}{x_1^n}) \end{aligned}$$

By using Definition II.1 (2), we have

$$=n|\frac{x}{x_1}|^{n-1}.\frac{1}{x_1}D(a_n\odot x_1^n,0)$$

Now we assume $G\in\mathbb{R}_{\mathcal{F}}$ be the largest of the fuzzy numbers $|a_1\odot x_1|,\ |a_2\odot x_1^2|,\ \ldots,|a_N\odot x_1^N|,1$. Then for all $n\in N\cup\{0\}$, we have $D(a_n\odot x_1^n,0)\leq G$, so

$$\leq n|\frac{x}{x_1}|^{n-1}.\frac{G}{x_1}$$

Therefore

$$\sum_{n=1}^{\infty}D(na_n\odot x^{n-1},0)\leq\frac{G}{|x_1|}\sum_{n=0}^{\infty}n|\frac{x}{x_1}|^{n-1}$$

By applying the ratio test for the real series $\sum_{n=0}^{\infty}n|\frac{x}{x_1}|^{n-1}$, it follows that the real series is absolutely convergent therefor by using real comparison test we have $|\frac{x}{x_1}|<1$. Then by using Definition III.12, we have the fuzzy series $\sum_{n=1}^{\infty}na_n\odot x^{n-1}$ is also fuzzy absolutely convergent Definition IV.1 and so according to Definition IV.1, is convergent. Now we just need to show that $\rho_{m_1}\ngeq\rho_m$, so suppose ρ_{m_1} is the radius of convergence of $\sum_{n=1}^{\infty}na_n\odot x^{n-1}$, $\rho_{m_1}>\rho_m$ and let $\rho_{m_1}>|x_2|>\rho_m$. If $|x_2|>\rho_m$ according Remark III.10 (ii), the fuzzy generalized power series $\sum_{n=0}^{\infty}a_n\odot x_2^n$, is diverges.

According to the first part of the present proof if $\rho_{m_1} > |x_2|$ it follows that fuzzy generalized power series $\sum_{n=0}^{\infty} na_n \odot x_2^n$ is absolutely convergent (Definition IV.1). So multiplying at $|x_2|$, we have

$$|x_2| \sum_{n=1}^{\infty} |na_n \odot x^{n-1}| = \sum_{n=1}^{\infty} |na_n \odot x^n|$$

By considering Lemma IV.1 (a), the fuzzy generalized power series $\sum_{n=1}^{\infty} |na_n \odot x^n|$ is convergent thus the convergence radius does not change. Let n is positive integer therefore

$$|a_n \odot x_2| \leq n|a_n \odot x_2| = |na_{n\odot} x_2|$$

Then

$$\sum_{n=1}^{\infty} |a_n \odot x_2| \leq n \sum_{n=1}^{\infty} |a_n \odot x_2| = \sum_{n=1}^{\infty} |na_n \odot x_2|$$

By applying the fuzzy comparison test (Theorem III.12) it follows the fuzzy generalized power series $\sum_{n=1}^{\infty} |a_n \odot x_2|$ is convergent, then $\sum_{n=0}^{\infty} |a_n \odot x_2|$ is convergent. Therefore, the hypothesis that $\rho_{m_1} > \rho_m$, is false then the proof is completed. □

Theorem IV.3. *Suppose that the fuzzy generalized power series $f(x) = \sum_{n=0}^{\infty} a_n \odot x^n$ is convergent for $x = x_0$ such that $x \neq 0$, then $f(x)$ is absolute convergent for all real values of x such that $|x| < |x_\ell|$.*

Proof. According to the Definition III.11 if a fuzzy generalized power series $\sum_{n=0}^{\infty} a_n \odot x_t^n$ is convergent then $\lim_{n\to\infty} a_n \odot x_i^n = 0$. And also by using Definition III.12 with assuming $\delta = 1$, we obtain

$$D(a_n \odot x^n, 0) < 1$$

We want to prove that $\sum_{n=0}^{\infty} a_n \odot x^n$ is also convergent, Since x is any number such that $|x| < |x_\ell|$, we get

$$\begin{aligned} D(a_n \odot x^n, 0) &= D(a_n \odot x_\ell^n \frac{x^n}{x_\ell^n}, 0.\frac{x_\ell^n}{x_\ell^n}) \\ &= |\frac{x^n}{x_\ell^n}| D(a_n \odot x_\ell^n, 0) \\ &\leq |\frac{x^n}{x_\ell^n}| \end{aligned}$$

So

$$D(a_n \odot x^n, 0) \leq |\frac{x^n}{x_\ell^n}| \tag{8}$$

$$D(\sum_{n=0}^{\infty} a_n \odot x^n, 0) \leq \sum_{n=0}^{\infty} |\frac{x^n}{x_\ell^n}|$$

By using comparison test for equation (8) and according Definition IV.1, then $\sum_{n=0}^{\infty} a_n \odot x^n$ is absolute convergent for $|x| < |x_\ell|$. □

Theorem IV.4. *Let for all $n = 0, 1, 2, ...$, $a_n \in \mathbb{R}_{\mathcal{F}}$, and $x \in \mathbb{R}$ if a fuzzy generalized power series $\sum_{n=0}^{\infty} a_n \odot x_v^n$ is divergent for x_v , then it is divergent for all values of x for which $|x| > |x_v|$.*

Proof. We assume with contradiction, that the fuzzy generalized power series $\sum_{n=0}^{\infty} a_n \odot x_v^n$ is convergent for some values of x, which $|x| > |x_v|$. According to Theorem IV.3 a fuzzy generalized power series $\sum_{n=0}^{\infty} a_n \odot x_v^n$ is convergent for x_v. So, this contradicts the hypothesis, then for all values of x for which $|x| > |x_v|$, the fuzzy generalized power series is divergent. □

V. Fuzzy Linear Inhomogeneous Differential Equation

Throughout this paper, we suppose that the fuzzy linear differential equation of second order of the form

$$\begin{cases} p(x) f''_{gH}(x) \oplus q(x) \odot f'_{gH}(x) \oplus r(x) \odot f(x) = 0, \\ f(x_0) = f_0 \in \mathbb{R}_{\mathcal{F}}, \\ f'(x_0) = f_1 \in \mathbb{R}_{\mathcal{F}}. \end{cases} \tag{9}$$

for which $x = x_0$ is an ordinary point, has the general solution $f_h : (-\rho_0, \rho_0) \to \mathbb{R}_{\mathcal{F}}$, where ρ_0 is a constant with $\rho_0 \in (0, \infty]$ and the coefficients $p, q, r : (-\rho_0, \rho_0) \to \mathbb{R}$, are analytic at 0 and have power series expansions $p(x) = \sum_{n=0}^{\infty} p_n x^n$, $q(x) = \sum_{n=0}^{\infty} q_n x^n$, $r(x) = \sum_{n=0}^{\infty} r_n x^n$, for all $x \in (-\rho_0, \rho_0)$. Since $x = 0$ is an ordinary point of Eq. (9), we remark that $p_0 \neq 0$. In this section, we are going to solve the fuzzy linear inhomogeneous differential equation of second order of the form

$$\begin{cases} p(x) f''_{gH}(x) \oplus q(x) \odot f'_{gH}(x) \oplus r(x) \odot f(x) \\ \quad = \sum_{n=0}^{\infty} a_n \odot (x)^n, \\ f(x_0) = f_0 \in \mathbb{R}_{\mathcal{F}}, \\ f'(x_0) = f_1 \in \mathbb{R}_{\mathcal{F}}. \end{cases} \tag{10}$$

under the assumption that $x = x_0 = 0$ is an ordinary point of the related fuzzy linear homogeneous differential equation (9) where ρ is radius of convergence and $p, q, r : (-\rho_0, \rho_0) \to \mathbb{R}$, are analytic at 0 such that for all $x \in (-\rho_0, \rho_0)$, $p(x), q(x)$ and $r(x)$ are real power series.

Theorem V.1. *Let $\sum_{n=0}^{\infty} a_n \odot x^n$ is fuzzy generalized power series of type (1) with radius of convergence $\rho_1 > 0$ and also assume that for all $n \in N \cup \{0\}$ there exists a fuzzy sequence $\{c_n\}$ satisfying the recurrence relation. If a_0, a_1 are arbitrary fuzzy positive constants and a_{n+2} (for $n \geq 0$) is defined by the following recursion formula*

$$\begin{aligned} a_n = \sum_{k=0}^{n} \Big[&(k+1)(k+2) \odot c_{k+2} \odot p_{n-k} \\ &\oplus (k+1) \odot c_{k+1} \odot q_{n-k} \oplus c_k \odot r_{n-k} \Big]. \end{aligned}$$

If ρ_2 is the radius of convergence of fuzzy generalized power series $\sum_{n=0}^{\infty} c_n \odot x^n$ and $\rho_3 = min\{\rho_0, \rho_1, \rho_2\}$, such that (ρ_0, ρ_0) is the domain of the general solution to (9)*. Then every fuzzy solution of the fuzzy linear inhomogeneous differential equation under $[(i) - gH]-$differentiability can be expressed by*

$$f(x) = f_h(x) \oplus \sum_{n=0}^{\infty} c_n \odot x^n.$$

for all $x \in (-\rho_3, \rho_3)$, where $f_h(x)$ is a fuzzy solution of (9)*.*

Proof. *According to the hypotheses of this theorem $x = 0$ is an ordinary point and $p(x)$, $q(x)$, $r(x)$ at the origin are real analytic functions, therefore, $T(x) = \sum_{n=0}^{\infty} p_n x^n = p_0 + p_1 x + p_2 x^2 + \cdots + p_n x^n + \cdots$, $q(x) = \sum_{n=0}^{\infty} q_n x^n = q_0 + q_1 x + q_2 x^2 + \cdots + q_n x^n + \cdots$ and $r(x) = \sum_{n=0}^{\infty} r_n x^n = r_0 + r_1 x + r_2 x^2 + \cdots + r_n x^n + \cdots$, by using to Theorems III.5, III.4, III.1and we substitute $c(x) = \sum_{n=0}^{\infty} c_n \odot x^n$ for $f(x), f'_{gH}(x)$ and $f''_{gH}(x)$ in* (10)*, we have*

$$p(x) \odot f''_{i-gH}(x) = \sum_{n=0}^{\infty} \Big[\sum_{k=0}^{n} (k+2)(k+1) p_{n-k} \odot c_{k+2} \Big] x^n,$$

$$q(x) \odot f'_{i-gH}(x) = \sum_{n=0}^{\infty} \Big[\sum_{k=0}^{n} (k+1) q_{n-k} \odot c_{k+1} \Big] x^n,$$

$$r(x) \odot f(x) = \sum_{n=0}^{\infty} \Big[\sum_{k=0}^{n} r_{n-k} \odot c_k \Big] x^n.$$

By substituting the above equations in Eq (10) *and adding fuzzy generalized power series terms and using Theorem III.3, then*

$$\begin{aligned} & p(x) f''_{gH}(x) \oplus q(x) \odot f'_{gH}(x) \oplus r(x) \odot f(x) \\ &= \sum_{n=0}^{\infty} \Big[\sum_{k=0}^{n} (k+2)(k+1) p_{n-k} \odot c_{k+2} \Big] x^n \\ &\oplus \sum_{n=0}^{\infty} \Big[\sum_{k=0}^{n} (k+1) q_{n-k} \odot c_{k+1} \Big] x^n \oplus \sum_{n=0}^{\infty} \Big[\sum_{k=0}^{n} r_{n-k} \odot c_k \Big] x^n \\ &= \sum_{n=0}^{\infty} \Big[\sum_{k=0}^{n} (k+2)(k+1) p_{n-k} \odot c_{k+2} \oplus (k+1) q_{n-k} \odot c_{k+1} \\ &= \oplus r_{n-k} \odot c_k \Big] x^n = \sum_{n=0}^{\infty} a_n \odot x^n. \end{aligned}$$

for all $x \in (-\rho_3, \rho_3)$. That is, $c(x) = \sum_{n=0}^{\infty} c_n \odot x^n$ is a particular solution of the fuzzy linear inhomogeneous differential equation (10)*, and hence every solution $f : (-\rho_3, \rho_3) \to \mathbb{R}_{\mathcal{F}}$ of equation* (10) *can be expressed by*

$$f(x) = f_h(x) \oplus \sum_{n=0}^{\infty} c_n \odot x^n.$$

where $f_h(x)$ is a fuzzy solution of the fuzzy linear homogeneous differential equation (9)*.* □

Theorem V.2. *Suppose that $\sum_{n=0}^{\infty} a_n \odot x^n$ is fuzzy generalized power series of type (2) with radius of convergence $\rho_1 > 0$ and also assume that for all $n \in N \cup \{0\}$ there exists a fuzzy sequence $\{c_n\}$ satisfying the recurrence relation*

$$\begin{aligned} a_n = \ominus \Big[& \sum_{k=0}^{n} (k+1)(k+2) \odot c_{k+2} \odot p_{n-k} \oplus (k+1) \\ & \odot c_{k+1} \odot q_{n-k} \oplus (-1)(c_0 \odot r_n) \oplus \sum_{k=1}^{n} c_k \odot r_{n-k} \Big]. \end{aligned}$$

If ρ_2 is the radius of convergence of fuzzy generalized power series $c_0 \ominus (-1) \sum_{n=0}^{\infty} c_n \odot x^n$ and $\rho_3 = min\{\rho_0, \rho_1, \rho_2\}$, such that (ρ_0, ρ_0) is the domain of the general solution to (9)*. Then every fuzzy solution of the fuzzy linear inhomogeneous differential equation under $[(i) - gH]-$differentiability can be expressed by*

$$f(x) = f_h(x) \oplus c_0 \ominus (-1) \sum_{n=1}^{\infty} c_n \odot x^n.$$

for all $x \in (-\rho_3, \rho_3)$, where $f_h(x)$ is a fuzzy solution of (9)*.*

Proof. *According to the hypotheses of this theorem $x = 0$ is an ordinary point and $p(x)$, $q(x)$, $r(x)$ at the origin are real analytic functions, therefore, $T(x) = \sum_{n=0}^{\infty} p_n x^n = p_0 + p_1 x + p_2 x^2 + \cdots + p_n x^n + \cdots$, $q(x) = \sum_{n=0}^{\infty} q_n x^n = q_0 + q_1 x + q_2 x^2 + \cdots + q_n x^n + \cdots$ and $r(x) = \sum_{n=0}^{\infty} r_n x^n = r_0 + r_1 x + r_2 x^2 + \cdots + r_n x^n + \cdots$, we substitute $c(x) = c_0 \ominus (-1) \sum_{n=0}^{\infty} c_n \odot x^n$ for $f(x), f'_{gH}(x)$ and $f''_{gH}(x)$ in* (10)*, so*

$$\begin{aligned} p(x) \odot f''_{i-gH}(x) &= \ominus(-1) \sum_{n=0}^{\infty} \Big[\sum_{k=0}^{n} (k+2)(k+1) p_{n-k} \\ &\quad \odot c_{k+2} \Big] x^n, \\ q(x) \odot f'_{i-gH}(x) &= \ominus(-1) \sum_{n=0}^{\infty} \Big[\sum_{k=0}^{n} (k+1) q_{n-k} \\ &\quad \odot c_{k+1} \Big] x^n, \\ r(x) \odot f(x) &= \sum_{n=0}^{\infty} r_n \odot c_0 \ominus (-1) \sum_{n=0}^{\infty} \\ &\quad \Big[\sum_{k=1}^{n} r_{n-k} \odot c_k \Big] x^n. \end{aligned}$$

By substituting the above equations in Eq (10) *and adding fuzzy generalized power series terms (Theorem III.3) then*

$$\begin{aligned} & p(x) f''_{gH}(x) \oplus q(x) \odot f'_{gH}(x) \oplus r(x) \odot f(x) \\ &= \ominus(-1) \sum_{n=0}^{\infty} \Big[\sum_{k=0}^{n} (k+2)(k+1) p_{n-k} \odot c_{k+2} \Big] x^n \\ &\ominus (-1) \sum_{n=0}^{\infty} \Big[\sum_{k=0}^{n} (k+1) q_{n-k} \odot c_{k+1} \Big] x^n \oplus \sum_{n=0}^{\infty} r_n \\ &\odot c_0 \ominus (-1) \sum_{n=0}^{\infty} \Big[\sum_{k=1}^{n} r_{n-k} \odot c_k \Big] x^n \\ &= \ominus(-1) \sum_{n=0}^{\infty} a_n \odot x^n. \end{aligned}$$

for all $x \in (-\rho_3, \rho_3)$. That is, $c(x) = c_0 \ominus (-1) \sum_{n=0}^{\infty} c_n \odot x^n$ is a particular solution of the fuzzy linear inhomogeneous differential equation (10)*, and hence every solution $f : (-\rho_3, \rho_3) \to \mathbb{R}_{\mathcal{F}}$ of equation* (10) *can be expressed by*

$$f(x) = f_h(x) \oplus c_0 \ominus (-1) \sum_{n=1}^{\infty} c_n \odot x^n.$$

where $f_h(x)$ is a fuzzy solution of the fuzzy linear homogeneous differential equation (9)*.* □

Theorem V.3. *Suppose that $\sum_{n=0}^{\infty} a_n \odot x^n$ is fuzzy generalized power series of type (3) with radius of convergence $\rho_1 > 0$*

and also assume that for all $n \in N \cup \{0\}$ there exists a fuzzy sequence $\{c_n\}$ satisfying the recurrence relation

$$\ominus^n (-1)^n a_n = \ominus \sum_{k=0}^{n} \Big[\ominus^{k+1} (-1)^{k+1} a_n (k+1)(k+2) \odot c_{k+2} \odot p_{n-k} \ominus (-1)(k+1) \odot c_{k+1} \odot q_{n-k} \ominus^k (-1)^k c_k \odot r_{n-k} \Big].$$

If ρ_2 is the radius of convergence of fuzzy generalized power series $\sum_{n=0}^{\infty} c_n \odot x^n$ and $\rho_3 = min\{\rho_0, \rho_1, \rho_2\}$, such that (ρ_0, ρ_0) is the domain of the general solution to (9)*. Then every fuzzy solution of the fuzzy linear inhomogeneous differential equation under $[(i) - gH]-$differentiability can be expressed by*

$$f(x) = f_h(x) \oplus \sum_{n=0}^{\infty} \ominus^n (-1)^n a_n \odot (x)^n,$$
$$\begin{cases} n = 2k, & \ominus^n (-1)^n \ a_n = \oplus \ a_n; \\ n = 2k+1, & \ominus^n (-1)^n \ a_n = \ominus (-1) \ a_n. \end{cases}$$

for all $x \in (-\rho_3, \rho_3)$, where $f_h(x)$ is a fuzzy solution of (9)*.*

Proof. *According to the hypotheses of this theorem $x = 0$ is an ordinary point and $p(x)$, $q(x)$, $r(x)$ at the origin are real analytic functions, therefore, $T(x) = \sum_{n=0}^{\infty} p_n x^n = p_0 + p_1 x + p_2 x^2 + \cdots + p_n x^n + \cdots$, $q(x) = \sum_{n=0}^{\infty} q_n x^n = q_0 + q_1 x + q_2 x^2 + \cdots + q_n x^n + \cdots$ and $r(x) = \sum_{n=0}^{\infty} r_n x^n = r_0 + r_1 x + r_2 x^2 + \cdots + r_n x^n + \cdots$, we substitute $c(x) = \ominus^n (-1)^n \sum_{n=0}^{\infty} c_n \odot x^n$ for $f(x), f'_{gH}(x)$ and $f''_{gH}(x)$ in* (10)*, so*

$$p(x) \odot f''_{i-gH}(x) = \sum_{n=0}^{\infty} \Big[\sum_{k=0}^{n} \ominus^k (-1)^k (k+2)(k+1) p_{n-k} \odot c_{k+2} \Big] x^n,$$
$$q(x) \odot f'_{i-gH}(x) = \sum_{n=0}^{\infty} \Big[\sum_{k=0}^{n} \ominus^{k+1} (-1)^{k+1} (k+1) q_{n-k} \odot c_{k+1} \Big] x^n,$$
$$r(x) \odot f(x) = \sum_{n=0}^{\infty} \Big[\sum_{k=0}^{n} \ominus^k (-1)^k r_{n-k} \odot c_k \Big] x^n.$$

By substituting the above equations in Eq (10) *and adding fuzzy generalized power series terms (Theorem III.3) then*

$$\begin{aligned} & p(x) f''_{gH}(x) \oplus q(x) \odot f'_{gH}(x) \oplus r(x) \odot f(x) \\ &= \sum_{n=0}^{\infty} \Big[\sum_{k=0}^{n} (k+2)(k+1) p_{n-k} \odot c_{k+2} \Big] x^n \\ &\oplus \sum_{n=0}^{\infty} \Big[\sum_{k=0}^{n} (k+1) q_{n-k} \odot c_{k+1} \Big] x^n \\ &\oplus \sum_{n=0}^{\infty} \Big[\sum_{k=0}^{n} \ominus^k (-1)^k r_{n-k} \odot c_k \Big] x^n \\ &= \ominus^n (-1)^n \sum_{n=0}^{\infty} \Big[\sum_{k=0}^{n} \ominus^{k+1} (-1)^{k+1} (k+2)(k+1) p_{n-k} \\ &\odot c_{k+2} \ominus^{k+1} (-1)^{k+1} (k+1) q_{n-k} \odot c_{k+1} \ominus^k (-1)^k r_{n-k} \\ &\odot c_k \Big] x^n = \ominus^k (-1)^k \sum_{n=0}^{\infty} a_n \odot x^n. \end{aligned}$$

for all $x \in (-\rho_3, \rho_3)$. That is, $c(x) = \ominus^n (-1)^n \sum_{n=0}^{\infty} c_n \odot x^n$ is a particular solution of the fuzzy linear inhomogeneous differential equation (10)*, and hence every solution $f : (-\rho_3, \rho_3) \to \mathbb{R}_{\mathcal{F}}$ of equation* (10) *can be expressed by*

$$f(x) = f_h(x) \ominus^n (-1)^n \sum_{n=0}^{\infty} c_n \odot x^n.$$

where $f_h(x)$ is a fuzzy solution of the fuzzy linear homogeneous differential equation (9)*.* □

Corollary 1. *Suppose that $p(x)$, $q(x)$ and $r(x)$ are real polynomials of degree at most $d \geq 0$ and d_0 be the degree of $p(x)$ and also radius of convergence of fuzzy generalized power series $\sum_{n=0}^{\infty} a_n \odot x^n$ and that there exists a fuzzy sequence $\{c_n\}$ satisfying the recurrence formula*

$$a_n = \sum_{k=n_0}^{n} \Big[(k+1)(k+2) \odot c_{k+2} \odot p_{n-k} \oplus (k+1) \odot c_{k+1} \odot q_{n-k} \oplus c_k \odot r_{n-k} \Big]. \tag{11}$$

for any $n \in N_0$, where $n_0 = max\{0, m-d\}$. If the fuzzy positive sequence $\{c_n\}$ satisfies the following conditions:

(i). $\lim_{n\to\infty} (c_{n-1} \div_g nc_n) = 0$,

(ii). *There exists a fuzzy number L such that $\lim_{n\to\infty} (c_n \div_g c_{n-1}) = L$ and $p_{d_0} \oplus L p_{d_0-1} \oplus ... \oplus L^{d_0-1} p_1 \oplus L^{d_0} p_0 \neq 0$.*

Then every fuzzy solution $f : (-\rho_3, \rho_3) \to \mathbb{R}^+_{\mathcal{F}}$ of the fuzzy linear inhomogeneous differential equation (10) *under $[(i) - gH]-$differentiability can be expressed by*

$$f(x) = f_h(x) \oplus \sum_{n=0}^{\infty} c_n \odot x^n. \tag{12}$$

for all $x \in (-\rho_3, \rho_3)$, where $\rho_3 = min[\rho_0, \rho_1]$ and $f_h(x)$ is a fuzzy solution of the fuzzy linear homogeneous differential equation (10)*.*

Proof. *Let n be any sufficiently large integer. According to the hypotheses of this theorem $p(x)$, $q(x)$ and $r(x)$ are real polynomials of degree at most $d \geq 0$ so $p_{d+1} = p_{d+2} = ... =$*

0, $q_{d+1} = q_{d+2} = ... = 0$ *and* $r_{d+1} = r_{d+2} = ... = 0$, *now we substitute* $m - d + k$ *for* k *in* (11)*, then we have*

$$a_n = \sum_{k=0}^{d} \Big[(n-d+k+2)(n-d+k+1) \odot c_{n-d+k+2} \odot p_{d-k} \oplus (n-d+k+1) \odot c_{n-d+k+1} \odot q_{d-k} \oplus c_{n-d+k} \odot r_{d-k}\Big].$$

By factoring and using condition (i) we have

$$\limsup_{n\to\infty} |a_n|^{\frac{1}{n}} = \limsup_{n\to\infty} \Big| \sum_{k=0}^{d} (n-d+k+2)(n-d+k+1) \odot c_{n-d+k+2} \odot \Big(p_{d-k} \oplus \frac{c_{n-d+k+1} \odot q_{d-k}}{(m-d+k+2)} \odot \frac{1}{c_{n-d+k+2}} \oplus \frac{c_{n-d+k} \odot r_{d-k}}{(n-d+k+2)(n-d+k+1)} \frac{c_{n-d+k+1}}{c_{n-d+k+1}} \odot \frac{1}{c_{n-d+k+2}}\Big)\Big|^{\frac{1}{n}}$$

$$= \limsup_{n\to\infty} \Big| \sum_{k=0}^{d} (n-d+k+2)(m-d+k+1) \odot c_{n-d+k+2} \odot p_{d-k}\Big|^{\frac{1}{n}}$$

Using condition (ii) we have

$$= \limsup_{n\to\infty} \Big| \sum_{k=d-d_0}^{d} (n-d+k+2)(m-d+k+1) \odot c_{n-d+k+2} \odot p_{d-k}\Big|^{\frac{1}{n}}$$

$$= \limsup_{n\to\infty} \Big| \sum_{k=d-d_0}^{d} (n-d_0+2)(n-d_0+1) \odot c_{n-d_0+2} \odot (p_{d_0} \oplus Lp_{d_0-1} \oplus ... \oplus L^{d_0-1}p_1 \oplus L^{d_0}p_0)\Big|^{\frac{1}{n}}$$

$$= \limsup_{n\to\infty} \Big|(n-d_0+2)(n-d_0+1)(p_{d_0} \oplus Lp_{d_0-1} \oplus ... \oplus L^{d_0-1}p_1 \oplus L^{d_0}p_0)\Big|^{\frac{1}{n}} \odot \Big(|c_{n-d_0+2}|^{\frac{1}{(n-d_0+2)}}\Big)^{\frac{(n-d_0+2)}{n}}$$

$$= \limsup_{n\to\infty} |c_{n-d_0+2}|^{\frac{1}{(n-d_0+2)}}.$$

which implies that the radius of convergence of the fuzzy generalized power series $\sum_{n=0}^{\infty} c_n \odot x^n$ *is* ρ_1. □

VI. Conclusion

The main purpose of this paper was to examine the existence and uniqueness of a sfuzzy Linear Inhomogeneous Differential equation under generalized Hukuhara differentiability with real coefficients. To achieve these results, first, some new concepts such as fuzzy absolute convergence, fuzzy analytic equations were studied. Then some important theorems such as fuzzy analytic functions, ordinary and singular points, were proved. Finally, the uniqueness of the fuzzy solution of fuzzy linear Inhomogeneous second-order differential equations in different scenarios was investigated and showed that the second-order fuzzy differential equations can have a unique solution in the form of fuzzy generalized power series by attention to the type of $gH-$differentiability.

References

[1] S. Abbasbandy, J.J. Nieto, M. Alavi,"Tuning of reachable set in one dimensional fuzzy differential inclusions," Chaos, Solitons and Fractals, vol. 26, pp. 1337–1341, December 2005.

[2] S. Abbasbandy, T. Allahviranloo, O. Lopez-Pouso, J.J. Nieto,"Numerical methods for fuzzy differential inclusions," Computers and mathematics with applications, vol. 48, pp. 1633–1641, November–December 2004.

[3] T. Allahviranloo, N. Ahmady, E. Ahmady, "Numerical solution of fuzzy differential equations by predictor-corrector method," Information Sciences, vol. 177, pp. 1633–1647, April 2007.

[4] T. Allahviranloo, Z. Gouyandeh, A. Armand, "A full fuzzy method for solving differential equation based on Taylor expansion," IOS Press, Intelligent and Fuzzy Systems, vol. 29, pp. 1039–1055, 2015.

[5] B. Bede, L. Stefanini, "Generalized differentiability of fuzzy-valued functions," Fuzzy Sets and Systems, vol. 230, pp. 119–141, November 2013.

[6] B. Bede , S. G. Gal , "Generalization of the differentiability of fuzzy-number-valued functions with applications to fuzzy differential equations," Fuzzy Sets and System, vol. 151, pp. 581–589, May 2005.

[7] S.L. Chang, L. A. Zadeh, "On fuzzy mapping and control, IEEE Trans, Systems Man Cybemet," vol. SMC-2, pp. 30–34, 1972.

[8] M. Chen, C. Wu, X. Xue, G. Liu, "On fuzzy boundary value problems, Information Sciences," vol. 178, pp. 1877–1892, April 2008.

[9] D. Dubios, H. Prade, "Towards fuzzy differential calculus: Part 3 differentiation," Fuzzy Sets and Systems, vol. 8, pp. 225–233, September 1982.

[10] M. Friedman, M. Ma, A. Kandel, "Numerical methods for calculating the fuzzy integral," Fuzzy Sets Systemn, vol. 83, pp. 57–62, October 1996.

[11] M. Hukuhara, " Integration des applications measurables dont lavaleur est un compact convexe," Funkcialaj Ekvacioj, vol. 10, pp. 205–223, 1967.

[12] O. Kaleva, "Fuzzy differential equations," Fuzzy Sets and Systems, vol. 24, pp. 301–317, December 1987.

[13] O. Kaleva, S. Seikkala, "On fuzzy metric spaces, Fuzzy Sets and Systems," vol. 12, pp. 215 –229, April 1984.

[14] U. Kadak, H. Efe, "On Uniform Convergence of Sequences and Series of Fuzzy-Valued Functions," Hindawi, Journal of function spaces, vol. 2015, pp. 1–11, September 2014.

[15] U. Kadak, F. Baar, "Power series of fuzzy numbers with real or fuzzy coefficients," Faculty of Sciences and

Mathematics, University of Nis Serbia, vol. 26, pp. 519–528, September 2012.

[16] M. Keshavarz, T. Allahviranloo, S. Abbasbandy, and M. H. Modarressi, “Analytical fuzzy solution of the ventricular pressure equation and prediction of the blood pressure,” (CMDE) Computational Methods for Differential Equations, Vol. 9, pp. 919–939, 2021.

[17] V. Lakshmikantham, T. Bhaskar , J. Devi, “Theory of set differential equations in metric spaces,” cambridge scientific publishers, 2006.

[18] R. Rodriguez-Lopez, “Comparison results for fuzzy differential equations Information Sciences,” Information Sciences, vol. 178, pp. 1756–1779, March 2008.

[19] Kh. Sabzi, T. Allahviranlooa, S. Abbasbandy, “A fuzzy generalized power series method under generalized Hukuhara differentiability for solving fuzzy Legendre diferential equation,” Soft Computing, vol 24,pp 8763-–8779, Jun 2020.

[20] S. Seikala, “On the fuzzy initial value problem, Fuzzy Sets and systems,” vol. 24, pp. 319–330, December 1987.

[21] L. Stefanini, “A generalization of Hukuhara difference and division for interval and fuzzy arithmetic,” Fuzzy sets system, vol. 161, pp. 1564–1584, June 2010.

[22] L. Stefanini, B. Bede, “Generalized Hukuhara differentiability of interval-valued functions and interval differential equations,” Nonlinear Analysis: Theory methods and applications, vol. 71, 209, pp. 1311–1328, February 2008.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Application Of Machine Learning Methods in Diagnosis of Alzheimer Disease Based on Fractal Feature Extraction and Convolutional Neural Network

Morteza Amini[1,2], Mir Mohsen Pedram[2,3]

[1]*Department of Cognitive Modeling, Shahid Beheshti University, Tehran, Iran,* mor_amini@sbu.ac.ir
[2]*Department of Cognitive Modeling, Institute for Cognitive Science Studies, Tehran, Iran*
[3]*Department of Electrical and Computer Engineering, Faculty of Engineering, Kharazmi University, Tehran, Iran*

***Abstract*— Alzheimer's disease impairs one's capacity to make sound strategic and operational decisions in authentic settings. It may be harder to adhere efficiently to ordinary concerns such as a stove that is burning or unexpected driving scenarios. Alzheimer's disease is characterized by memory loss. The inability to recall recent events or discussions is one of the first indicators. Memory problems intensify as the disease develops, and additional symptoms emerge. This research presents a CNN-based architecture with specific processes starting with image acquisition and ending with AD-classification to categorize scanned MRI images in order to predict whether or not they contain Alzheimer's disease utilizing a machine learning application and digital image processing. Six machine learning algorithms for identifying and segmenting MRI scans of patients' brains are presented in this paper. Six machine learning approaches were used to identify and segment ultrasound pictures in order to determine the illness or tumor kind. First, the fractal approach is used to extract the picture characteristics. The photos of the individuals were then classified using KNN, SVM, DT, and NB classification algorithms. The suggested high-potential CNN methodology may be utilized to identify AD using MRI scans, according to the results of this study.**

***Index Terms* — Alzheimer's disease, Deep learning, Classification, Diagnosis.**

I. INTRODUCTION

Alzheimer's disease (AD) is a neurological illness that worsens with time. It's a neurological condition in which brain cells die, resulting in dementia and cognitive deterioration. In addition, it is the most frequent kind of dementia [1,2]. It has a profoundly damaging influence on people's personal and social lives. According to recent figures, more than 46.8 million individuals in the United States today have dementia, with 44 million having Alzheimer's disease. By 2050, the population will have risen to 131.5 million [3,4]. Mild Cognitive Impairment (MCI) is a stage between Cognitive and emotional Healthy and Dementia, with a 10% conversion efficiency to Alzheimer's disease [5].

It is critical to diagnose the condition early in order to intervene. Clinical trials in Alzheimer's disease tend to include participants at earlier stages of neuronal deterioration, when therapy is frequently more successful [6]. In this respect, scanning is one of the most attractive study avenues for early identification of Alzheimer's disease, as the increasing deterioration of brain structures may be detected in structural MRI scans as a significant cerebral shrinkage (MRI). In scientific science, imaging techniques have been developed. They are undertaking technical changes to evaluate diagnostic pictures and give a unified view of the diagnosis for physicians and researchers, as well as academics, to promote further analyzation [7]. Digital radiography, mammograms, ultrasonography, MRI, computed tomography equipment, and other diagnostic imaging modal approaches survive today that are used to create clinical pictures [8]. There has been a lot of study towards obtaining brain pictures from MRI databases that have been linked to the gathering of magnetic resonance. The fact that MRI scanned pictures provide better spatial qualities is a considerable benefit. For the sake of illness diagnosis, the visual features are significant [9]–[12]. Alzheimer's disease (AD) is a neurological illness that affects people as they become older, causing loss of memory impairment. With the aging population, its frequency is expected to rise in the next two decades, providing a significant burden to society. Neuron loss is the most noticeable feature of AD pathology, accompanied by brain shrinkage that progresses from AD characteristic areas

(e.g., the hippocampus and amygdala) to the whole cortical plate, which may be seen with an MRI scan [13].

Various ml algorithms applied to structural MRI have been utilized in previous research to classify Alzheimer's patients. The most widely used data mining approach is the Support Vector Machine [14]. To develop the classifier that simplify the AD detection, SVM extracts high-dimensional, significant attributes from MRI. Feature extraction, feature selection, dimension reduction, and feature-based classifications suggests that technological are the four processes of a machine learning categorization research study [15]. This method has a number of limitations, including the requirement for complicated picture pre-processing, a long processing time, and intensive calculations. Furthermore, the repeatability of these procedures is seen as a difficulty.

According to a recent analysis, a significant proportion of studies utilizing convolutional neural networks (CNNs) for Alzheimer's disease either have obvious data leaking issues or provide a brief explanation for the validation procedure to guarantee that data leakage has not happened. The potential for data leaking simply emphasizes the importance of evaluating a separate group of photographs. This review is in addition to, Pinaya et al.[16] calculated how much each patient differed from the norm and identified which brain areas were involved. They tested the normative precision to that of conventional detectors. For Alzheimer disease categorization, Ashraf [17] used several CNN-based transfer learning approaches. On the benchmark ADNI dataset, they used several parameters and obtained exceptional accuracy. On the ADNI dataset, they examined 13 various types of pre-trained neural network models employing a fine-tuned transfer learning strategy across two domains (94 AD, 138 MCI and 146 NC). DenseNet, on the other hand, performed better, obtaining a maximum average accuracy of 99.05. Lee et al.[18] presented a deep CNN data iteration strategy for utilizing MRI to classify Alzheimer's disease. Slice selection was offered as a way to accomplish AlexNet's benefits. The results of their experiments revealed that their data permutation strategy enhanced accuracy of classification for AD classification. On the ADNI database, the overall accuracy for binary and ternary classification are 98.74 percent and 98.06 percent, accordingly.

This research presents a CNN-based methodology with specific processes starting with image acquisition and ending with AD-classification to categorize scanned Region of interest in order to predict whether or not they contain Alzheimer's disease utilizing a machine learning application and image analysis.

II. METHODS AND MATERIALS

This research presents a CNN-based paradigm for AD classification.

By using MRI image volume as data, we suggest a novel CNN for the multiclass AD biomarker detection challenge in this paper. The five primary processes in our pipeline are: brain extraction and data filtering, feature extraction, standardization, CNN computation, and domain setup (see Figure 1). This section describes our workflow in depth, including picture pre-processing, CNN designs, and dynamic programming.

Figure 1. The process diagram of the presented approach

Learning discriminative patterns invariant to many transformations is required to optimize deep-learning systems employing MRIs in their original space, which necessitates new designs and an even greater number of data, with all predicted changes [19]. We may anticipate comparable components to be generally in the same spatial location if we register our photographs to a standard template, allowing us to process the full image at once and auto detect the most essential parts of interest. To retrieve and normalize brain pictures, we employed the normalization tools. We used the default options, such as transformation types, sequence, and metrics that were given. The following processes made up our brain extraction and normalizing pipeline: picture intensity, bias field correction, translations realignment, stiff convert, affine transform, flexible symmetric standardization, and deployment of brain masking from the map.

Because we were working with registered brains, we chose a map that was less rigid and linear, allowing for some variance during the application procedure. This map also had a depth information, ensuring that finer features were preserved. The generated picture has the same size (i.e., 256 × 256) as the atlas after the brain separation and standardization process. We deleted the boundary dimensions that had no data, resulting in a final image, since the brain is encompassed in a smaller region inside the image. This extra area accounts for 48% of the initial mass, lowering sparsity from 78 to 54%. Finally, we computed mean and variance on the training set, then utilized these to normalize the data to zero mean and unit variance across all sets. We employed a single-pass online mean and estimated variance approach because the used datasets could not fit in main memory. The biggest problem we faced was putting up a registration process, which included the chosen atlas, that delivered contextually relevant results in a fair amount of time while still maintaining quality.

A domain adaption technique was investigated in addition to neural impulses and CNN processors. In our strategy, we trained a system using one dataset and then tested it on a second dataset (i.e., ADNI). Despite their similarities, such discrepancies indicate that the origin histogram and the target data distribution may diverge. Thus, even if just a limited number of samples from this target domain are used, it should be feasible to enhance the results by modifying the previously-trained upgrading to a new Database. This example is more closely comparable to a real-world scenario and is known as cross dataset validation. Throughout training and real use, data sets will most typically vary. Furthermore, this is a more trustworthy approach of evaluating a machine learning algorithm’s capability. In one of the last CNN layers, we used our improved CNN to features extracted from the whole target database, which was the beginning of our topic effective approach. Then, using just the

target training set to derive the parameters, we standardized these variables to zero mean. On the whole target training set, we optimized a one-vs.-rest logistic regression using the normalized data. We utilized grid search with leave-one-out cross-validation to discover the optimal parameters for this classifier. Then, eventually, we had a solution that was optimized for the source domain, allowing us to provide classifier chances for each sample in the scope. This pipeline resembles a transfer-learning strategy.

III. Results and Discussion

A. Data collection

Our data gathering approach at ADNI originally contained affect participation communities: CN, MCI, and AD. The MCI stage was known as early MCI (eMCI) and late MCI starting in ADNI. Later on, a category with neurocognitive complaints was established. We used only similar diagnoses since one of our key aims for this study was to do well in the ADNI task [20,21]. As a result, the eMCI and lMCI phases were combined with MCI, but SMC was ignored. We obtained all accessible raw T1-weighted MRI images related with Alzheimer's disease from these sources, ignoring any pre- or post-processed images [22,23].

B. Feature selection

Model categorization is based on fractal characteristics collected from MRI scans. The histogram of the pictures on the photos is recovered using the fractal approach, as illustrated in Figure 2.

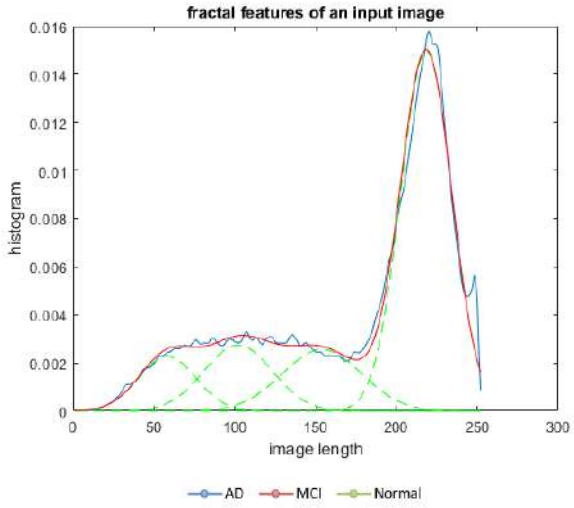


Figure 2. A histogram of characteristics retrieved from MRI scans

The photos are translated to a histogram and analyzed using the fractal approach, as seen in Figure 2. As a consequence, four charts of the standard normal value replace the derived model. As elements of each picture, the properties of the resulting distributions are produced in the form of four integers. The characteristics have been recorded in a matrix and are now ready to be categorized. The picture histogram's blue line is shown in Figure 2. We picked four characteristics with greater accuracy by rearranging the photos used for modeling. The summation of the functions is shown by a red line. Gaussian plotted are image characteristics. This procedure is followed for all photos in the data set. As a result, utilizing four features, the categorization data set is transformed to a matrix.

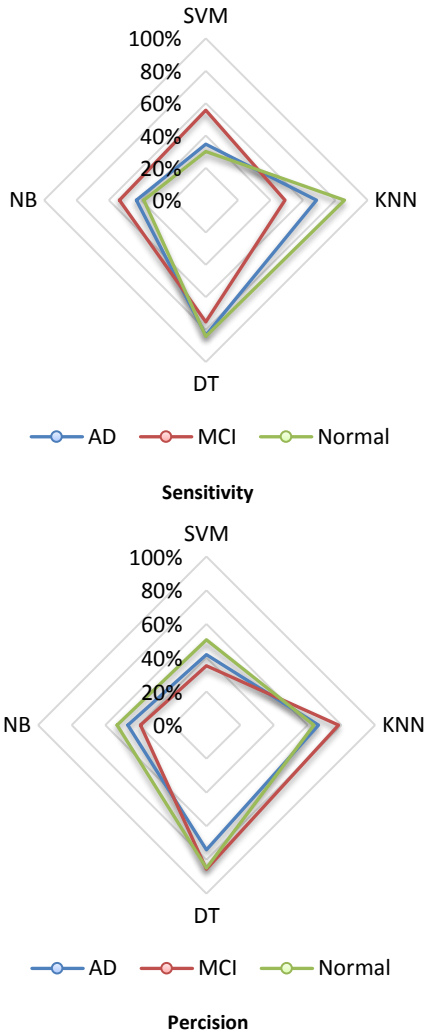


Figure 3. The confusion matrices for determining tumor kind and illness

We produced training, verification, and test splits for the Dataset. We divided randomly the related individuals in Dataset, retaining the original age, sex, and clinical categorization across each set, and using 70% of the subjects for training, 10% for validation, and 20% for testing. We allocated photos from prior subjects to the relevant collection in each Dataset after that. Only fresh participants were considered in the stratified random split.

Alzheimer's disease is identified and diagnosed utilizing fractal aspects of pictures in this section of the study. The fractal approach is one of the most effective ways, as is feature extraction in pictures, particularly MRI; tumors have been detected using this method and well-known categories. For each image, the feature extraction result is four scalars, which are utilized as classification inputs. Furthermore, all classes' outer layers were designated 0 for normal, 1 for AD, and 2 for MCI. The suggested models are intended to identify different types of tumors. Confusion matrices are used to visualize the categorization results.

Four traditional classifiers have been chosen to recognize the kind of tumor, as per the figure, and are among the most powerful approaches for identification and therapy. For basic detection, these categories frequently produce excellent results. There are a lot of issues with various classes (for example, triplets in this case). According to the results of the decision tree technique in Figure 3, a reasonable quantity of information can aid in the diagnosis of a tumor. In this procedure, 122 photos (84.2%) were accurately classified out of 133 photographs with normal. In addition, 83.5 percent of individuals with Alzheimer's disease were accurately diagnosed from 111 photographs. There were 133 photos in this category of schooling that were accurately recognized. With Alzheimer's disease, 10 photos were classified as healthy and 12 as MCI, with 22 erroneous images reported. Eventually, the decision tree method's accuracy was 81 percent, with a 19 percent error rate. SVM and NB were reported with 67.7%, 40.1 percent, and 44.9 percent accuracy, accordingly, after KNN. This degree of precision was insufficient to meet the classification's requirements. As a result, we need to create a model that can more reliably and sensitively identify the condition. As a result, the suggested model based on a convolutional neural network is discussed in the next chapter.

C. Classification of Alzheimer's disease using the given CNN approach

The outcomes of CNN architecture in Alzheimer's disease are provided in this section. There are three categories in this classification: AD pictures, MCI, and healthy or ordinary images. The CNN architecture model is trained with database photos, and the following evaluation criteria are used to examine the model:

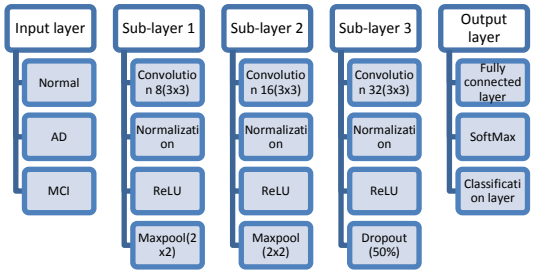


Figure 4. CNN's framework for identifying or identifying Alzheimer's disease kind.

The framework of CNN's suggested approaches for identifying cancers is shown in Figure 4. This network has 16 layers and three convolutional layers. Patients with an AD of one have an MCI of two, while healthy people have an MCI of zero. Furthermore, 70% of the picture samples are used for training stage and 30% for model validation. The findings are shown in section below. The amount of loss and accuracy as a proportion of training epochs are shown in Figure 5. During network training, these diagrams are shown. After 7000 rounds, the process came to a conclusion with the best outcome with more accuracy and fewer network loss. The cross-validation of the procedure is shown by the back points.

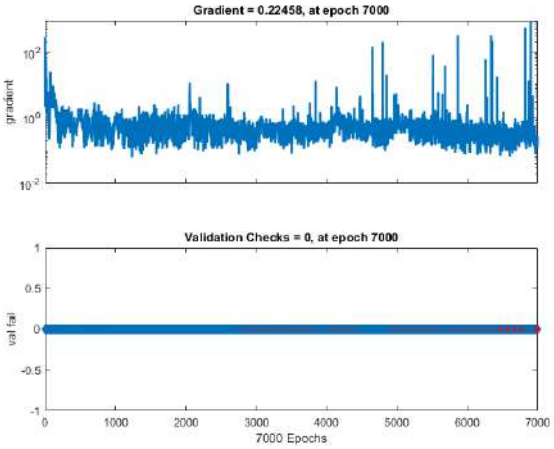


Figure 5. The Clinical reasoning model's accuracy and loss throughout program execution.

Both the training and trial sets are used to assess the final trained model. This prediction may be seen in Figure 6 of the confusion matrices. The model containing 306 photos of AD correctly learned 305 items (99.7%) in the training dataset, as seen in this figure. In addition, 100 photos out of 147 MCI and 93 healthy tissue images are accurately forecast. As a consequence, the training set's accuracy is 99.8 percent. The experimenter, which represents 30% of the raw document and did not participate in the analysis, produced the following findings.

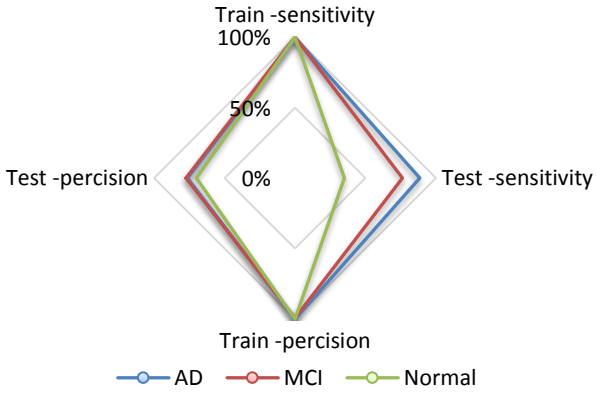


Figure 6. The Clinical reasoning model's accuracy and loss throughout program execution.

According to the test findings, 116 of the 131 AD experimental pictures were successfully identified. To put it another way, this diagnosis had an 88.5 percent sensitivity. In addition, 48 of the 63 MCI photos (76.2 percent) were accurately identified, with 15 images (23.8 percent) misinterpreted as AD, resulting in erroneous findings. This model was shown to have a low sensitivity of 35% for identifying healthy tissues that were not included in the analysis. To put it another way, the model's overall accuracy for the validation of the proposed model is 76.1 percent. Table 1 shows the results of comparing the proposed methods for tumor classification. According to Figure 7, the CNN model correctly detected tumors with a substantially

greater accuracy than previous approaches, resulting in a considerable improvement. The area under the ROC curve (AUC) is another measure of the success of classification models, which obtained 96 percent for the suggested model. The suggested high-potential CNN algorithm can be utilized to diagnose Alzheimer's Disease MRI pictures, according to the finding of the study.

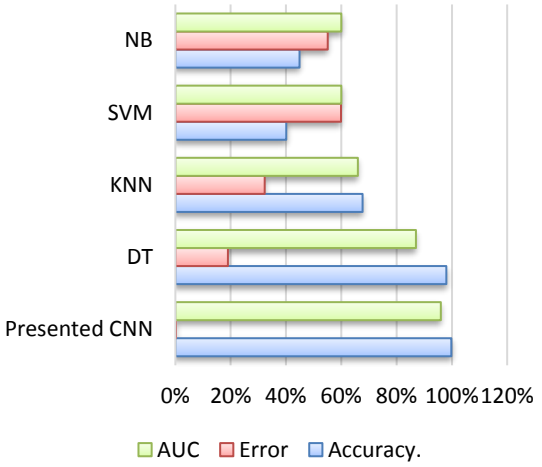


Figure 7. Comparison of different classification models.

IV. CONCLUSION

We constructed a CNN using the brain picture as input using ADNI statistics. With a network design based on the proposed CNN approach, we were able to attain the highest accuracy. Our method is totally automated (i.e., no further data input or operator interaction is required) and extremely quick. Notably, unlike all of the other submissions, our solution did not rely on any domain-specific data from Alzheimer's disease. As a result, we believe it might be used to treat additional illnesses that could benefit from the CAD system's use of MRI data as input. We believe that our method may be used to detect significant patterns in data, confirm earlier expert conclusions, aid in diagnosis situations, and ultimately benefit recognize trends for disorders other than Alzheimer's disease. Our results are backed up by our explainable artificial intelligence approaches. The accuracy of the decision tree approach is 81 percent, while the error rate is 19 percent, according to the data. SVM and NB were reported with 67.7%, 40.1 percent, and 44.9 percent accuracy, respectively, after KNN. The suggested CNN method's final framework was tested once it has been trained in both the training and experimental sets. In 306 photos with AD, the given model properly trained 305 examples (99.7%). As a consequence, the training set's accuracy is 99.8 percent. In terms of the test findings, 116 out of 131 AD experimental photos were accurately identified, implying that this diagnosis had an 88.5 percent sensitivity. To put it another way, the model's overall accuracy for the evaluation of proposed concept is 76.1 percent. The suggested high-potential CNN algorithm may be utilized to diagnose AD using MRI scans, according to the outcome of the study.

References

[1] I. Bhushan, M. Kour, G. Kour, S. Gupta, S. Sharma, and A. Yadav, "Alzheimer's disease: Causes \& treatment--A review," Ann Biotechnol, vol. 1, no. 1, p. 1002, 2018.

[2] G. J. Biessels and F. Despa, "Cognitive decline and dementia in diabetes mellitus: mechanisms and clinical implications," Nat. Rev. Endocrinol., vol. 14, no. 10, pp. 591–604, 2018, doi: 10.1038/s41574-018-0048-7.

[3] P. Wanleenuwat, P. Iwanowski, and W. Kozubski, "Alzheimer's dementia: pathogenesis and impact of cardiovascular risk factors on cognitive decline," Postgrad. Med., vol. 131, no. 7, pp. 415–422, 2019, doi: 10.1080/00325481.2019.1657776.

[4] Amini M, Pedram MM, Moradi A, Ochani M. Highly Sensitive Surface Plasmon Resonance Biosensors Utilizing Prism-Waveguide Configuration for Detection of Alzheimer Disease Biomarker. Plasmonics. 2021 Sep 8:1-8.

[5] K. Moradi, S. Faghani, A. Abdolalizadeh, M. Khomeijani-Farahani, and A. Ashraf-Ganjouei, "Biological Features of Reversion from Mild Cognitive Impairment to Normal Cognition: A Study of Cerebrospinal Fluid Markers and Brain Volume," J. Alzheimer's Dis. Reports, vol. 5, no. 1, pp. 179–186, 2021, doi: 10.3233/ADR-200229.

[6] G. Folego, M. Weiler, R. F. Casseb, R. Pires, and A. Rocha, "Alzheimer's Disease Detection Through Whole-Brain 3D-CNN MRI," Front. Bioeng. Biotechnol., vol. 8, 2020, doi: 10.3389/fbioe.2020.534592.

[7] K. R. Kruthika, Rajeswari, and H. D. Maheshappa, "CBIR system using Capsule Networks and 3D CNN for Alzheimer's disease diagnosis," Informatics Med. Unlocked, vol. 14, pp. 59–68, 2019, doi: 10.1016/j.imu.2018.12.001.

[8] F. Pesapane, M. Codari, and F. Sardanelli, "Artificial intelligence in medical imaging: threat or opportunity? Radiologists again at the forefront of innovation in medicine," Eur. Radiol. Exp., vol. 2, no. 1, pp. 1–10, 2018, doi: 10.1186/s41747-018-0061-6.

[9] R. Ferri et al., "Stacked autoencoders as new models for an accurate Alzheimer's disease classification support using resting-state EEG and MRI measurements," Clin. Neurophysiol., vol. 132, no. 1, pp. 232–245, 2021, doi: 10.1016/j.clinph.2020.09.015.

[10] R. Reda, A. Zanza, A. Mazzoni, A. Cicconetti, L. Testarelli, and D. Di Nardo, "An update of the possible applications of magnetic resonance imaging (Mri) in dentistry: A literature review," J. Imaging, vol. 7, no. 5, p. 75, 2021, doi: 10.3390/jimaging7050075.

[11] E. A. S. El-Dahshan, H. M. Mohsen, K. Revett, and A. B. M. Salem, "Computer-aided diagnosis of human brain tumor through MRI: A survey and a new algorithm," Expert Syst. Appl., vol. 41, no. 11, pp. 5526–5545, 2014, doi: 10.1016/j.eswa.2014.01.021.

[12] R. M. Berquist et al., "The digital fish library: Using MRI to digitize, database, and document the morphological diversity of fish," PLoS One, vol. 7, no. 4, p. e34499, 2012, doi: 10.1371/journal.pone.0034499.

[13] P. A. Boyle and D. A. Bennett, "Scope of dementia: Epidemiology and public health impact," in Long-Term Management of Dementia, CRC Press, 2016, pp. 1–24.

[14] S. Lahmiri and A. Shmuel, "Performance of machine learning methods applied to structural MRI and ADAS cognitive scores in diagnosing Alzheimer's disease," Biomed. Signal Process. Control, vol. 52, pp. 414–419, 2019, doi: 10.1016/j.bspc.2018.08.009.

[15] I. Beheshti and H. Demirel, "Probability distribution function-based classification of structural MRI for the detection of Alzheimer's disease," Comput. Biol. Med., vol. 64, pp. 208–216, 2015, doi: 10.1016/j.compbiomed.2015.07.006.

[16] W. H. L. Pinaya et al., "Using normative modelling to detect disease progression in mild cognitive impairment and Alzheimer's disease in a cross-sectional multi-cohort study," Sci. Rep., vol. 11, no. 1, pp. 1–13, 2021, doi: 10.1038/s41598-021-95098-0.

[17] A. Ashraf, S. Naz, S. H. Shirazi, I. Razzak, and M. Parsad, "Deep transfer learning for alzheimer neurological disorder detection," Multimed. Tools Appl., vol. 80, no. 20, pp. 30117–30142, 2021, doi: 10.1007/s11042-020-10331-8.

[18] B. Lee, W. Ellahi, and J. Y. Choi, "Using deep CNN with data permutation scheme for classification of Alzheimer's disease in structural

magnetic resonance imaging (SMRI)," IEICE Trans. Inf. Syst., vol. E102D, no. 7, pp. 1384–1395, 2019, doi: 10.1587/transinf.2018EDP7393.

[19] A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI," Z. Med. Phys., vol. 29, no. 2, pp. 102–127, 2019, doi: 10.1016/j.zemedi.2018.11.002.

[20] Amini M, Pedram MM, Moradi A, Ochani M. Highly Sensitive Surface Plasmon Resonance Biosensors Utilizing Prism-Waveguide Configuration for Detection of Alzheimer Disease Biomarker. Plasmonics. 2021 Sep 8:1-8.

[21] Ouchani M, Gharibzadeh S, Jamshidi M, Amini M. A Review of Methods of Diagnosis and Complexity Analysis of Alzheimer's Disease Using EEG Signals. BioMed Research International. 2021 Oct 27;2021.

[22] Amini M, Moradi A, Jamshidi M, Ouchani M. Single and combined neuroimaging techniques for Alzheimer's disease detection. Computational Intelligence and Neuroscience. 2021 Jul 13;2021.

[23] Amini M, Pedram M, Moradi A, Ouchani M. Diagnosis of Alzheimer's Disease by Time-Dependent Power Spectrum Descriptors and Convolutional Neural Network Using EEG Signal. Computational and Mathematical Methods in Medicine. 2021 Apr 24;2021.

[24] Amini M, Pedram M, Moradi A, Ouchani M. Diagnosis of Alzheimer's Disease Severity with fMRI Images Using Robust Multitask Feature Extraction Method and Convolutional Neural Network (CNN). Computational and Mathematical Methods in Medicine. 2021 Apr 27;2021.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Approximation of derivation–homomorphism fuzzy functional inequalities in matrix valued FC-$\diamond$-algebras

1st Safoura Rezaei Aderyani, 2nd Reza Saadati

1st , School of Mathematics, Iran University of Science and Technology, Tehran, Iran, safoura_rezaei99@mathdep.iust.ac.ir
2nd , School of Mathematics, Iran University of Science and Technology, Tehran, Iran, rsaadati@iust.ac.ir

***Abstract*— We introduce and solve two different additive-additive $(\alpha,\beta)-$fuzzy functional inequalities. In addition, we investigate derivations and homomorphisms on matrix valued FB-algebras and unital matrix valued FC-$\diamond$-algebras, associated to the additive-additive $(\alpha,\beta)-$fuzzy functional inequalities. Finally, as an application, we introduce new fuzzy control functions.**

***Index Terms*— Derivation and homomorphism, matrix valued FC-$\diamond$-algebras, additive-additive (α,β)-fuzzy functional inequalities.**

I. Introduction

In this paper, we introduce and solve the following additive-additive (α,β)-fuzzy functional inequality

$$\begin{aligned}&\Phi(\mu(j,a+b)-\mu(j,a)-\mu(j,b),\overrightarrow{t}) \qquad (1)\\ &\circledast\Phi(2\eta(j,\frac{a+b}{2})-\eta(j,a)-\eta(j,b),\overrightarrow{t})\\ &\succeq\Phi(\alpha[2\mu(j,\frac{a+b}{2})-\mu(j,a)-\mu(j,b)],\overrightarrow{t})\\ &\circledast\Phi(\beta[\eta(j,a+b)-\eta(j,a)-\eta(j,b)],\overrightarrow{t}),\end{aligned}$$

where $0\neq\alpha,\beta\in\mathbb{C}$ and $\max\{|\alpha|,|\beta|\}<1$. In addition, we study homomorphisms and derivations in matrix valued FB-algebras and unital matrix valued FC-$\diamond$-algebras, associated to the additive-additive (α,β)-fuzzy functional inequality (1).

Furthermore, we introduce and solve the following additive-additive (α,β)-fuzzy functional inequality

$$\begin{aligned}&\Phi(\mu(j,a+b+c)-\mu(j,a)-\mu(j,b)-\mu(j,c),\overrightarrow{t}) \qquad (2)\\ &\circledast\Phi(3\eta(j,\frac{a+b+c}{3})\\ &\quad+\eta(j,a-2b+c)+\eta(j,a+b-2c)-3\eta(j,a),\overrightarrow{t})\\ &\succeq\Phi(\alpha[3\mu(j,\frac{a+b+c}{3})-\mu(j,a)-\mu(j,b)-\mu(j,c)],\overrightarrow{t})\\ &\quad\circledast\Phi(\beta[\eta(j,a+b+c)+\eta(j,a-2b+c)\\ &\qquad+\eta(j,a+b-2c)-3\eta(j,a)],\overrightarrow{t}),\end{aligned}$$

where $0\neq\alpha,\beta\in\mathbb{C}$ and $\max\{|\alpha|,|\beta|\}<1$. In addition, we study FC-$\diamond$-ternary derivations and FC-$\diamond$-ternary homomorphisms in matrix valued FC-$\diamond$-ternary algebras, associated to the additive-additive (α,β)-fuzzy functional inequality (2).

II. Preliminaries

Here, we let $\mathfrak{C}=[0,1]$. Let

$$\begin{aligned}\text{diag}\mathsf{H}_n(\mathfrak{C})=\Bigg\{&\begin{bmatrix}\mathscr{H}_1&&\\&\ddots&\\&&\mathscr{H}_n\end{bmatrix}\\ &=\text{diag}[\mathscr{H}_1,\cdots,\mathscr{H}_n],\ \mathscr{H}_1,...,\mathscr{H}_n\in\mathfrak{C}\Bigg\}.\end{aligned}$$

We denote $\mathscr{H}:=\text{diag}[\mathscr{H}_1,\cdots,\mathscr{H}_n]\preceq\mathscr{G}:=\text{diag}[\mathscr{G}_1,\cdots,\mathscr{G}_n]$ if and only if $\mathscr{H}_i\leq\mathscr{G}_i$ for any $1\leq i\leq n$, and note that $\mathbf{0}=\begin{bmatrix}0&&\\&\ddots&\\&&0\end{bmatrix}$ and $\mathbf{1}=\begin{bmatrix}1&&\\&\ddots&\\&&1\end{bmatrix}$.

Now, we define GTN on $\text{diag}\mathsf{H}_n(\mathfrak{C})$.

Definition 1. *A GTN on $diag\mathsf{H}_n(\mathfrak{C})$ is an operation $\circledast: diag\mathsf{H}_n(\mathfrak{C})\times diag\mathsf{H}_n(\mathfrak{C})\to diag\mathsf{H}_n(\mathfrak{C})$ satisfying the following conditions:*

- $(\forall\mathscr{H}\in diag\mathsf{H}_n(\mathfrak{C}))(\mathscr{H}\circledast\mathbf{1})=\mathscr{H})$;
- $(\forall(\mathscr{H},\mathscr{G})\in(diag\mathsf{H}_n(\mathfrak{C}))^2)(\mathscr{H}\circledast\mathscr{G}=\mathscr{G}\circledast\mathscr{H})$;
- $(\forall(\mathscr{H},\mathscr{G},\mathscr{K})\in(diag\mathsf{H}_n(\mathfrak{C})^3)(\mathscr{H}\circledast(\mathscr{G}\circledast\mathscr{K})=(\mathscr{H}\circledast\mathscr{G})\circledast\mathscr{K})$;
- $(\forall(\mathscr{H},\mathscr{H}',\mathscr{G},\mathscr{G}')\in(diag\mathsf{H}_n(\mathfrak{C}^4)(\mathscr{H}\preceq\mathscr{H}'$ *and* $\mathscr{G}\preceq\mathscr{G}'\Longrightarrow\mathscr{H}\circledast\mathscr{G}\preceq\mathscr{H}'\circledast\mathscr{G}'$.

For every $\mathscr{H},\mathscr{G}\in\text{diag}\mathsf{H}_n(\mathfrak{C})$ and every sequences $\{\mathscr{H}_k\}$ and $\{\mathscr{G}_k\}$ converging to $\mathscr{H}$ and $\mathscr{G}$ suppose we have

$$\lim_k(\mathscr{H}_k\circledast\mathscr{G}_k)=\mathscr{H}\circledast\mathscr{G},$$

then, $\circledast$ on $\text{diag}\mathsf{H}_n(\mathfrak{C})$ is continuous (in short, CGTN). Consider the following example of CGTN.

Let $\circledast_M : \text{diag}\mathsf{H}_n(\mathfrak{C}) \times \text{diag}\mathsf{H}_n(\mathfrak{C}) \to \text{diag}\mathsf{H}_n(\mathfrak{C})$, such that,

$$\begin{aligned}\mathscr{H} \circledast_M \mathscr{G} &= \text{diag}[\mathscr{H}_1, \cdots, \mathscr{H}_n] \circledast_M \text{diag}[\mathscr{G}_1, \cdots, \mathscr{G}_n] \\ &= \text{diag}[\min\{\mathscr{H}_1, \mathscr{G}_1\}, \cdots, \min\{\mathscr{H}_n, \mathscr{G}_n\}],\end{aligned}$$

thus $\circledast_M$ is CGTN.

Suppose $\mathcal{W}$ is a vector space and $\overrightarrow{t} \in (0,\infty)^n$ for $n \in \mathbb{N}$, $\overrightarrow{t} = (t_1, ..., t_n)$, in which $t_j \in (0,\infty)$ for all $j = 1, ..., n$. Note that $\overrightarrow{0} \prec \overrightarrow{t}$ if and only if $0 < t_j, \forall j = 1, ..., n$ and $\overrightarrow{t} \longrightarrow \infty$ is equivalent to $t_j \longrightarrow \infty$, for all $j = 1, ..., n$.

We denote the set of matrix valued fuzzy set (MVF-set) by Ψ. Now, $\psi \in \Psi$ means that $\psi : \mathcal{W} \times (0,\infty)^n \to \text{diag}\mathsf{H}_n(\mathfrak{C})$ satisfies the following conditions:

- ψ is continuous;
- $\psi(\omega, .)$ is non-decreasing, where $\omega \in \mathcal{W}$;
- $\lim\limits_{\overrightarrow{\zeta} \longrightarrow +\infty} \psi(\omega, \overrightarrow{t}) = \mathbf{1}$, where $\omega \in \mathcal{W}$.

In Ψ we define $\preceq$ as follows

$$\psi \precsim \chi \iff \psi(\omega, \overrightarrow{t}) \preceq \chi(\omega, \overrightarrow{t'}),$$

for any $\omega \in \mathcal{W}$ and $\overrightarrow{t}, \overrightarrow{t'} \in (0,\infty)^n$.

Definition 2. *Consider the CGTN $\circledast$, a vector space $\mathcal{W}$ and MVF-set $\Phi : \mathcal{W} \times (0,\infty)^n \to diag\mathcal{M}_n(\mathfrak{C})$. In this case, we define a matrix valued fuzzy normed space (matrix valued FN-space) $(\mathcal{W}, \Phi, \circledast)$ as,*

- $\Phi(\omega, \overrightarrow{t}) = \mathbf{1}$, *for all* $\overrightarrow{t} \succ \overrightarrow{0}$ *if and only if* $w = 0$;
- $\Phi(\hbar\omega, \overrightarrow{t}) = \Phi(\omega, \dfrac{\overrightarrow{t}}{|\hbar|})$ *for all* $\omega \in \mathcal{W}$, $\overrightarrow{t} \succ \overrightarrow{0}$ *and* $\hbar \in \mathbb{C}$ *with* $\hbar \neq 0$;
- $\Phi(\omega+\omega', \overrightarrow{t} + \overrightarrow{t'}) \succeq \Phi(\omega, \overrightarrow{t}) \circledast \Phi(w', \overrightarrow{t'})$ *for all* $\omega, \omega' \in \mathcal{W}$ *and* $\overrightarrow{t}$, $\overrightarrow{t'} \succ \overrightarrow{0}$.
- $\lim\limits_{\overrightarrow{t} \longrightarrow +\infty} \Phi(\omega, \overrightarrow{t}) = \mathbf{1}$, *for all* $\omega \in \mathcal{W}$.

Definition 3. *Consider matrix valued FN-space $(\mathcal{W}, \Phi, \circledast)$ and the CGTN's $\circledcirc$, $\circledast$. If*

- $\Phi(ww', tt') \succeq \Phi(w,t) \circledcirc \Phi(w',t')$ *for any* $w, w' \in \mathcal{W}$ *and any* $t', t > 0$,

$(\mathcal{W}, \Phi, \circledcirc, \circledast)$ is called a matrix valued FN-algebra (matrix valued fuzzy normed algebra).

A complete matrix valued FN-algebra is called matrix valued fuzzy Banach algebra (or matrix valued FB-algebra). For more details you can see [1].

Consider the matrix valued FN-spaces $\mathscr{X}_1$ and $\mathscr{X}_2$. Let (J, Π, Φ) be the probability measure space. Let the Borel measureable spaces $(\mathscr{X}_1, \mathfrak{B}_{\mathscr{X}_1})$ and $(\mathscr{X}_2, \mathfrak{B}_{\mathscr{X}_2})$. a map $\mathcal{Q} : J \times \mathscr{X}_1 \to \mathscr{X}_2$ is a fuzzy operator if $\{j : \mathcal{Q}(j, \mathfrak{X}_1) \in \mathcal{C}\} \in \Pi$ for all $\mathfrak{X}_1$ in $\mathscr{X}_1$ and $\mathcal{C} \in \mathfrak{B}_{\mathscr{X}_2}$. $\mathcal{Q}$ is linear if

$$\mathcal{Q}(j, \lambda_1\mathfrak{X}_1 + \lambda_2\mathfrak{X}_2) = d\mathcal{Q}(j, \mathfrak{X}_1) + b\mathcal{Q}(j, \mathfrak{X}_2),$$

for any $\mathfrak{X}_1, \mathfrak{X}_2 \in \mathscr{X}_1$ and any $\lambda_1, \lambda_2 \in \mathbb{R}$. and also $\mathcal{Q}$ is bounded if we be able to find a $\mathsf{Q}(j) > 0$ so that

$$\Phi(\mathcal{Q}(j, \mathfrak{X}_1) - \mathcal{Q}(j, \mathfrak{X}_2), \mathsf{Q}(j)\overrightarrow{t}) \succeq \Phi(\mathfrak{X}_1 - \mathfrak{X}_2, \overrightarrow{t}),$$

for any $\mathfrak{X}_1, \mathfrak{X}_2 \in \mathscr{X}_1$ and any $\overrightarrow{t} \in (0,\infty)^n$.

Definition 4. *Suppose $(\Omega, \Phi, \circledcirc, \circledast)$ is matrix valued FB-algebra. An involution on Ω is a fuzzy operator $\rho \to \rho^\diamond$ from $J \times \Omega$ to Ω satisfying the following conditions for any $\overrightarrow{t} \in (0,\infty)^n$,*

- $\Phi(\rho^{\diamond\diamond}, \overrightarrow{t}) = \Phi(\rho, \overrightarrow{t})$ *for each* $\rho \in \Omega$.
- $\Phi((\sigma\rho + \sigma'\rho')^\diamond, \overrightarrow{t}) = \Phi(\overline{\sigma}\rho^\diamond + \overline{\sigma'}(\rho')^\diamond, \overrightarrow{t})$ *for each* $\rho \in \Omega$.
- $\Phi((\rho\rho')^\diamond, \overrightarrow{t}) = \Phi((\rho')^\diamond\rho^\diamond, \overrightarrow{t})$ *for each* $\rho, \rho' \in \Omega$.

Then Ω is called a matix valued FB-$\diamond$-algebra. In addition, if $\Phi(\rho^\diamond\rho, \overrightarrow{t}) = \Phi(\rho, \overrightarrow{t})$ for any $\rho \in \Omega$ and $\overrightarrow{t} \in (0,\infty)^n$, then Ω is called a matix valued FC-$\diamond$-algebra.

Now, we consider the unital matrix valued FC-$\diamond$-algebra $(\Omega, \Phi, \circledcirc, \circledast)$ with unite ϱ and unitary $U(\varrho) = \{\Bbbk \in \Omega : \Bbbk^\diamond\Bbbk = \Bbbk\Bbbk^\diamond = \varrho\}$.

Suppose Ω is a matrix valued FB-algebra. A $\mathbb{C}-$linear fuzzy operator $\mu : J \times \Omega \to \Omega$ is a derivation if $\mu(j, ab) = \mu(j, a)b + a\mu(j, b)$, for any $a, b \in \Omega$, and a $\mathbb{C}-$linear fuzzy operator $\eta : J \times \Omega \to \Omega$ is a homomorphism if $\eta(j, ab) = \eta(j, a)\eta(j, b)$, for any $a, b \in \Omega$.

A matrix valued FC-$\diamond$-ternary algebra is a matrix valued FB-space Ω, equipped with a ternary product $(a, b, c) \mapsto [a, b, c]$ of Ω^3 into Ω, which is $\mathbb{C}-$linear in the outer variables, conjugate $\mathbb{C}-$linear in the middle variable, and associative in the sense that $[a, b, [c, w, v]] = [a, [w, c, b], v] = [[a, b, c], w, v]$, and satisfies $\|[a, b, c]\| \leq \|a\|.\|b\|.\|c\|$ and $\|[a, a, a]\| = \|a\|^3$.

Suppose Ω is a matrix valued FC-$\diamond$-ternary algebra. A $\mathbb{C}-$linear fuzzy operator $\mu : J \times \Omega \to \Omega$ is a FC-$\diamond$-ternary derivation if $\mu(j, [a, b, c]) = [\mu(j, a), b, c] + [a, \mu(j, b), c] + [a, b, \mu(j, c)]$, for any $a, b, c \in \Omega$, and a $\mathbb{C}-$linear fuzzy operator $\eta : J \times \Omega \to \Omega$ is a FC-$\diamond$-ternary homomorphism if $\eta(j, [a, b, c]) = [\eta(j, a), \eta(, jb), (c)]$. For more details, you can see [2].

III. Main results

In this section Using fixed point alternative, we investigate aditive-aditive (α, β)-fuzzy functional inequalities (1) and (2), under some extra condition. For more details, you can see [2]–[11].

Throughout the paper, we let $\circledast_M = \circledast$.

A. Additive-additive (α, β)-fuzzy functional inequality (1) in in matrix valued FB-algebras

Here, suppose $(\Omega, \Phi, \circledast, \circledast)$ is matrix valued FB-algebra.

Lemma 1. *If fuzzy operators $\mu, \eta : J \times \Omega \to \Omega$ satisfy (1) and $\mu(j, 0) = \eta(j, 0) = 0$, then the fuzzy operators μ, η are additive.*

Proof: Letting $a = b$ in (1),

$$\Phi(\mu(j, 2a) - 2\mu(j, a), \overrightarrow{t}) \succeq \Phi(\beta[\eta(j, 2a) - 2\eta(j, a)], \overrightarrow{t}).$$

Now, let $b = 0$ in (1),

$$\Phi(2\eta(j, \frac{a}{2}) - \eta(j, a), \overrightarrow{t}) \succeq \Phi(\alpha[2\mu(j, \frac{a}{2}) - \mu(j, a)], \overrightarrow{t}),$$

and so

$$\Phi(2\eta(j, a) - \eta(j, 2a), \overrightarrow{t}) \succeq \Phi(\alpha[2\mu(j, a) - \mu(j, 2a)], \overrightarrow{t}),$$

for any $\overrightarrow{t} \in (0,\infty)^n$. Thus

$$\Phi(\mu(j,2a)-2\mu(j,a),\overrightarrow{t}) \succeq \Phi(\beta\alpha[\mu(j,2a)-2\mu(j,a)],\overrightarrow{t}).$$

$$\Phi(2\eta(j,a)-\eta(j,2a),\overrightarrow{t}) \succeq \Phi(\alpha\beta[2\eta(j,a)-\eta(j,2a)],\overrightarrow{t}),$$

since $\max\{|\alpha|,|\beta|\}<1$, therefore, $\eta(j,2a)=2\eta(j,a)$ and $\mu(j,2a)=2\mu(j,a)$. Now, according to (1),

$$\begin{aligned}&\Phi(\mu(j,a+b)-\mu(j,a)-\mu(j,b),\overrightarrow{t})\\ &\circledast\Phi(\eta(j,a+b)-\eta(j,a)-\eta(j,b),\overrightarrow{t})\\ &\quad\succeq \Phi(\alpha[\mu(j,a+b)-\mu(j,a)-\mu(j,b)],\overrightarrow{t})\\ &\quad\circledast\Phi(\beta[\eta(j,a+b)-\eta(j,a)-\eta(j,b)],\overrightarrow{t}),\end{aligned}$$

for any $\overrightarrow{t} \in (0,\infty)^n$. Thus, $\mu(j,a+b)-\mu(j,a)-\mu(j,b)=0$ and $\eta(j,a+b)-\eta(j,a)-\eta(j,b)=0$. ∎

Lemma 2. *Suppose $\gamma : J\times\Omega\to\Omega$ be a fuzzy operator such that*

$$\gamma(j,\lambda(a+b))=\lambda\gamma(j,a)+\lambda\gamma(j,b),$$

for any $\lambda\in\Xi:=\{\rho\in\mathbb{C}:|\rho|=1\}$. Then γ is $\mathbb{C}-$linear.

Theorem 1. *Suppose $\phi:\Omega^2\times(0,\infty)^n\to$ diag$\mathsf{H}_n(\mathfrak{C})$ is a control fuzzy function such that*

$$\lim_{k\to\infty}\phi(\frac{a}{2^k},\frac{b}{2^k},\frac{\overrightarrow{t}}{2^k})=\mathbf{1}. \tag{3}$$

Suppose $\mu,\eta : J\times\Omega\to\Omega$ are fuzzy operators satisfying $\mu(j,0)=\eta(j,0)=0$ and

$$\begin{aligned}&\Phi(\mu(j,\lambda(a+b))-\lambda\mu(j,a)-\lambda\mu(j,b),\overrightarrow{t}) \qquad (4)\\ &\circledast\Phi(2\eta(j,\lambda\frac{a+b}{2})-\lambda\eta(j,a)-\lambda\eta(j,b),\overrightarrow{t})\\ &\quad\succeq\Phi(\alpha[2\mu(j,\lambda\frac{a+b}{2})-\lambda\mu(j,a)-\lambda\mu(j,b)],\overrightarrow{t})\\ &\quad\circledast\Phi(\beta[\eta(j,\lambda(a+b))-\lambda\eta(j,a)-\lambda\eta(j,b)],\overrightarrow{t}),\end{aligned}$$

for any $\lambda\in\Xi$ and $\overrightarrow{t}\in(0,\infty)^n$. If μ,η satisfy

$$\begin{aligned}&\Phi(\mu(j,ab)-\mu(j,a)b-a\mu(j,b),\overrightarrow{t}) \qquad (5)\\ &\quad\circledast\Phi(\eta(j,ab)-\eta(j,a)\eta(j,b),\overrightarrow{t})\succeq\phi(a,b,\overrightarrow{t}),\end{aligned}$$

for any $\overrightarrow{t}\in(0,\infty)^n$, then the fuzzy operator μ is a derivation and the fuzzy operator η is a homomorphism.

Proof: Let $\lambda=1$, in (4). According to Lemma 1, η and μ are additive. It follows from (4) that

$$\begin{aligned}&\Phi(\mu(j,\lambda(a+b))-\lambda\mu(j,a)-\lambda\mu(j,b),\overrightarrow{t}) \qquad (6)\\ &\circledast\Phi(\eta(j,\lambda(a+b))-\lambda\eta(j,a)-\lambda\eta(j,b),\overrightarrow{t})\\ &\quad\succeq\Phi(\alpha[\mu(j,\lambda(a+b))-\lambda\mu(j,a)-\lambda\mu(j,b)],\overrightarrow{t})\\ &\quad\circledast\Phi(\beta[\eta(j,\lambda(a+b))-\lambda\eta(j,a)-\lambda\eta(j,b)],\overrightarrow{t}),\end{aligned}$$

for any $\overrightarrow{t}\in(0,\infty)^n,\lambda\in\Xi$, and $a,b\in\Omega$. Since $\max\{|\alpha|,|\beta|\}<1$, therefore $\mu(j,\lambda(a+b))-\lambda\mu(j,a)-\lambda\mu(j,b)=0$, and $\eta(j,\lambda(a+b))-\lambda\eta(j,a)-\lambda\eta(j,b)=0$, for any $\overrightarrow{t}\in(0,\infty)^n,\lambda\in\Xi$, and $a,b\in\Omega$. Using Lemma 2, η and μ are $\mathbb{C}-$linear.

According to (5), and the additivity of μ and η,

$$\begin{aligned}&\Phi(\mu(j,ab)-\mu(j,a)b-a\mu(j,b),\overrightarrow{t})\\ &\quad\circledast\Phi(\eta(j,ab)-\eta(j,a)\eta(j,b),\overrightarrow{t})\\ &=\Phi(2^{2n}[\mu(j,\frac{ab}{2^{2n}})-\mu(j,\frac{a}{2^n})\frac{b}{2^n}-\frac{a}{2^n}\mu(j,\frac{a}{2^n})]\overrightarrow{t})\\ &\quad\circledast\Phi(2^{2n}[\eta(j,\frac{ab}{2^{2n}})-\eta(j,\frac{a}{2^n})\eta(j,\frac{b}{2^n})]\overrightarrow{t})\\ &\succeq\phi(\frac{a}{2^n},\frac{b}{2^n},\frac{\overrightarrow{t}}{2^n}),\end{aligned}$$

for any $\overrightarrow{t}\in(0,\infty)^n$.

Since $\lim_{k\to\infty}\phi(\frac{a}{2^k},\frac{b}{2^k},\frac{\overrightarrow{t}}{2^k})=\mathbf{1}$, then $\mu(j,ab)-\mu(j,a)b-a\mu(j,b)=0$ and $\eta(j,ab)=\eta(j,a)\eta(j,b)$. Therefore, η is a homomorphism and μ is a derivation. ∎

Theorem 2. *Suppose $\phi:\Omega^2\times(0,\infty)^n\to$ diag$\mathsf{H}_n(\mathfrak{C})$ is a control fuzzy function such that*

$$\lim_{k\to\infty}\phi(2^ka,2^kb,2^k\overrightarrow{t})=\mathbf{1}. \tag{7}$$

Suppose $\mu,\eta : J\times\Omega\to\Omega$ are fuzzy operators satisfying $\mu(j,0)=\eta(j,0)=0$, (4) and (5) for any $\overrightarrow{t}\in(0,\infty)^n$. then μ is a derivation and η is a homomorphism.

Proof: Via the same reasoning presented in the proof of Theorem 2, we can show that μ and η are $\mathbb{C}-$linear. According to (5) and the additivity of η and μ,

$$\begin{aligned}&\Phi(\mu(j,ab)-\mu(j,a)b-a\mu(j,b),\overrightarrow{t})\\ &\quad\circledast\Phi(\eta(j,ab)-\eta(j,a)\eta(j,b),\overrightarrow{t})\\ &=\Phi(\frac{1}{2^{2n}}[\mu(j,2^{2n}ab)-\mu(j,2^na)2^nb-2^na\mu(j,2^nb)]\overrightarrow{t})\\ &\quad\circledast\Phi(\frac{1}{2^{2n}}[\eta(j,2^{2n}ab)-\eta(j,2^na)\eta(j,2^nb)]\overrightarrow{t})\\ &\succeq\phi(2^na,2^nb,2^n\overrightarrow{t}),\end{aligned}$$

for any $\overrightarrow{t}\in(0,\infty)^n$. Which tend $\mathbf{1}$ as $n\longrightarrow\infty$. So $\mu(j,ab)-\mu(j,a)b-a\mu(j,b)=0$ and $\eta(j,ab)=\eta(j,a)\eta(j,b)$. Therefore, η is a homomorphism and μ is a derivation. ∎

Theorem 3. *Suppose $\phi:\Omega^2\times(0,\infty)^n\to$ diag$\mathsf{H}_n(\mathfrak{C})$ is a control fuzzy function such that there exists a $\varpi<1$, with*

$$\phi(\frac{a}{2},\frac{b}{2},\overrightarrow{t})\succeq\phi(a,b,\frac{2\overrightarrow{t}}{\varpi}). \tag{8}$$

Suppose $\mu,\eta : J\times\Omega\to\Omega$ are fuzzy operators satisfying $\mu(j,0)=\eta(j,0)=0$, (4) and (5) for any $\overrightarrow{t}\in(0,\infty)^n$. then μ is a derivation and η is a homomorphism.

Proof: It follows from (8),

$$\begin{aligned}\lim_{n\to\infty}\phi(\frac{a}{2^n},\frac{b}{2^n},\frac{\overrightarrow{t}}{2^n}) &\succeq \phi(a,b,\frac{\overrightarrow{t}}{\frac{1}{2}\sum_{n=0}^{\infty}\varpi^n})\\ &\succeq \prod_{n=1}^{\infty}\phi(a,b,\frac{\overrightarrow{t}}{\frac{1}{2}\varpi^n})=\mathbf{1}.\end{aligned}$$

∎

Theorem 4. *Suppose $\phi : \Omega^2 \times (0,\infty)^n \to diag\mathsf{H}_n(\mathfrak{C})$ is a control fuzzy function such that there exists a $\varpi < 1$, with*

$$\phi(a,b,\overrightarrow{t}) \succeq \phi(\frac{a}{2}\frac{b}{2},\frac{\overrightarrow{t}}{2\varpi}). \tag{9}$$

Suppose $\mu,\eta : J \times \Omega \to \Omega$ are fuzzy operators satisfying $\mu(j,0) = \eta(j,0) = 0$, (4) and (5) for any $\overrightarrow{t} \in (0,\infty)^n$. then μ is a derivation and η is a homomorphism.

Proof: It follows from (9),

$$\begin{aligned}\lim_{n\to\infty}\phi(2^n a,2^n b,2^n\overrightarrow{t}) &\succeq \phi(a,b,\frac{\overrightarrow{t}}{\frac{1}{2}\sum_{n=0}^{\infty}\varpi^n})\\ &\succeq \prod_{n=1}^{\infty}\phi(a,b,\frac{\overrightarrow{t}}{\frac{1}{2}\varpi^n}) = \mathbf{1}.\end{aligned}$$

■

Remark 1. *In (5), the pair (η,μ) of a homomorphism η and a derivation μ can be replaced by the pair of a homomorphism and a homomorphism or the pair of a derivation and a derivation.*

B. Additive-additive (α,β)-fuzzy functional inequality (1) in matrix valued FC-$\diamond$-algebra

Throughout this section, let $(\Omega,\Phi,\circledast,\circledast)$ is a unital matrix valued FC-$\diamond$-algebra with unitary group $U(\Omega)$.

Theorem 5. *Suppose $\phi : \Omega^2 \times (0,\infty)^n \to diag\mathsf{H}_n(\mathfrak{C})$ is a control fuzzy function satisfying (3). Suppose $\mu,\eta : J\times\Omega \to \Omega$ are fuzzy operators satisfying $\mu(j,0) = \eta(j,0) = 0$ and (4). If μ,η satisfy*

$$\begin{aligned}&\Phi(\mu(j,\zeta\xi) - \mu(j,\zeta)\xi - \zeta\mu(j,\xi)),\overrightarrow{t})\\ &\quad\circledast\Phi(\eta(j,\zeta\xi) - \eta(j,\zeta)\eta(j,\xi),\overrightarrow{t})\\ &\succeq \phi(\zeta,\xi,\overrightarrow{t}),\end{aligned} \tag{10}$$

for any $\zeta,\xi \in U(\Omega)$, then η and μ are a homomorphism and derivation, respectively.

Proof: Suppose $\lambda = 1$, in (4). According to Lemma 1, η and μ are additive. It follows from (4),

$$\begin{aligned}&\Phi(\mu(j,\lambda(a+b)) - \lambda\mu(j,a) - \lambda\mu(j,b),\overrightarrow{t})\\ &\circledast\Phi(\eta(j,a+b) - \lambda\eta(j,a) - \lambda\eta(j,b),\overrightarrow{t})\\ &\quad\succeq \Phi(\alpha[\mu(j,a+b) - \lambda\mu(j,a) - \lambda\mu(j,b)],\overrightarrow{t})\\ &\quad\circledast\Phi(\beta[\eta(j,\lambda(a+b)) - \lambda\eta(j,a) - \lambda\eta(j,b)],\overrightarrow{t}),\end{aligned} \tag{11}$$

for any $\lambda \in \Xi$, and $\overrightarrow{t} \in (0,\infty)^n$. Since $\max\{|\alpha|,|\beta|\} < 1$, we have $\mu(j,\lambda(a+b)) - \lambda\mu(j,a) - \lambda\mu(j,b) = 0$, and $\eta(j,\lambda(a+b)) - \lambda\eta(j,a) - \lambda\eta(j,b) = 0$, for any $\lambda \in \Xi$, and $\overrightarrow{t} \in (0,\infty)^n$. Hence, by Lemma 2, η and μ are $\mathbb{C}-$linear. By a similar method to the proof of Theorem 3.1 in [2], the proof will be completed. ■

Theorem 6. *Suppose $\phi : \Omega^2 \times (0,\infty)^n \to diag\mathsf{H}_n(\mathfrak{C})$ is a fuzzy control function satisfying (7) and $\mu,\eta : J\times\Omega \to \Omega$ are fuzzy operators satisfying $\mu(j,0) = \eta(j,0) = 0$ and (4) and (10). Then η and μ are a homomorphism and a derivation, respectively.*

Proof: The proof is similar to the proofs of Theorem 2 and 5. ■

Remark 2. *through the fixed point method given in Theorem 3 and 4, we can get the same results in Theorem 5 and 6.*

C. Additive-additive (α,β)–fuzzy functional inequality (2) in matrix valued FC-$\diamond$-ternary algebra

Here, suppose $(\mho,\Phi,\circledast,\circledast)$ is matrix valued FC-$\diamond$-ternary algebra.

Lemma 3. *If fuzzy operators $\mu,\eta : J\times\mho \to \mho$ satisfy (2) and $\mu(j,0) = \eta(j,0) = 0$, then μ and η are additive.*

Proof: Let $a = b = c$ in (2). Thus, we have

$$\Phi(\mu(j,3a) - 3\mu(j,a),\overrightarrow{t}) \succeq \Phi(\beta[\eta(j,3a) - 3\eta(j,a)],\overrightarrow{t}),$$

for any $\overrightarrow{t} \in (0,\infty)^n$. Now, let $b = z = 0$ in (1), therefore

$$\Phi(3\eta(j,\frac{a}{3}) - 3\eta(j,a),\overrightarrow{t}) \succeq \Phi(\alpha[3\mu(j,\frac{a}{3}) - \mu(j,a)],\overrightarrow{t}),$$

and so

$$\Phi(3\eta(j,a) - \eta(j,3a),\overrightarrow{t}) \succeq \Phi(\alpha[3\mu(j,a) - \mu(j,3a)],\overrightarrow{t}),$$

for any $\overrightarrow{t} \in (0,\infty)^n$. Thus

$$\Phi(\mu(j,3a) - 3\mu(j,a),\overrightarrow{t}) \succeq \Phi(\alpha\beta[\mu(j,3a) - 3\mu(j,a)],\overrightarrow{t}),$$

$$\Phi(3\eta(j,a) - \eta(j,3a),\overrightarrow{t}) \succeq \Phi(\alpha\beta[3\eta(j,a) - \eta(j,3a)],\overrightarrow{t}),$$

for any $\overrightarrow{t} \in (0,\infty)^n$. Since $\max\{|a|,|b|\} < 1$, then, $\eta(j,3x) = 3\eta(j,x)$ and $\mu(j,3x) = 3\mu(j,x)$ for any $a \in \mho$. According to (1),

$$\begin{aligned}&\Phi(\mu(j,a+b+c) - \mu(j,a) - \mu(j,b) - \mu(j,c),\overrightarrow{t})\\ &\circledast\Phi(\eta(j,a+b+c) + \eta(j,a-2b+c)\\ &\quad+\eta(j,a+b-2c) - 3\eta(j,a),\overrightarrow{t})\\ &\succeq \Phi(\alpha[\mu(j,a+b+c) - \mu(j,a) - \mu(j,b) - \mu(j,c)],\overrightarrow{t})\\ &\quad\circledast\Phi(\beta[\eta(j,a+b+c) + \eta(j,a-2b+c)\\ &\qquad+\eta(j,a+b-2c) - 3\eta(j,a)],\overrightarrow{t}),\end{aligned}$$

for any $\overrightarrow{t} \in (0,\infty)^n$. Since $\max\{|a|,|b|\} < 1$, then, $\mu(j,a+b+c) - \mu(j,a) - \mu(j,b) - \mu(j,c) = 0$, and $\eta(j,a+b+c) + \eta(j,a-2b+c) + \eta(j,a+b-2c) - 3\eta(j,a) = 0$, for any $\overrightarrow{t} \in (0,\infty)^n$. So η and μ are additive. ■

Theorem 7. *Suppose $\phi : \mho^3 \times (0,\infty)^n \to diag\mathsf{H}_n(\mathfrak{C})$ is a control fuzzy function such that*

$$\lim_{k\to\infty}\phi(\frac{a}{2^k},\frac{b}{2^k},\frac{c}{2^k},\frac{\overrightarrow{t}}{2^k}) = \mathbf{1}. \tag{12}$$

Suppose $\mu,\eta : J\times\mho \to \mho$ are fuzzy operators satisfying $\mu(j,0) = \eta(j,0) = 0$ and

$$\begin{aligned}&\Phi(\mu(j,\lambda(a+b+c)) - \lambda\mu(j,a) - \lambda\mu(j,b) - \lambda\mu(j,c),\overrightarrow{t})\\ &\circledast\Phi(3\eta(j,\lambda\frac{a+b+c}{3})\\ &\quad+\lambda\eta(j,a-2b+c) + \lambda\eta(j,a+b-2c) - 3\lambda\eta(j,a),\overrightarrow{t})\\ &\succeq \Phi(\alpha[3\mu(j,\lambda\frac{a+b+c}{3}) - \lambda\mu(j,a) - \lambda\mu(j,b) - \lambda\mu(j,c)],\overrightarrow{t})\\ &\quad\circledast\Phi(\beta[\eta(j,\lambda(a+b+c)) + \lambda\eta(j,a-2b+c)\\ &\qquad+\lambda\eta(j,a+b-2c) - 3\lambda\eta(j,a)],\overrightarrow{t}),\end{aligned} \tag{13}$$

for any $\lambda \in \Xi$ and any $\overrightarrow{t} \in (0,\infty)^n$. If μ and η satisfy

$$\begin{aligned}&\Phi(\mu(j,[a,b,c]) - [\mu(j,a),b,c] - [a,\mu(j,b),c]\\ &-[a,b,\mu(j,c)],\overrightarrow{t}) \circledast \Phi(\eta(j,[a,b,c])\\ &-[\eta(j,a),\eta(j,b),\eta(j,c)],\overrightarrow{t})\\ &\succeq \phi(a,b,c,\overrightarrow{t}),\end{aligned} \tag{14}$$

for any $\overrightarrow{t} \in (0,\infty)^n$, then η is a FC-$\diamond$-ternary homomorphism and μ is a FC-$\diamond$-ternary derivation.

Proof: Let $\lambda = 1$, in (13). According to Lemma 3, η and μ are additive. Now, according to (13),

$$\begin{aligned}&\Phi(\mu(j,\lambda(a+b+c)) - \lambda\mu(j,a) - \lambda\mu(j,b) - \lambda\mu(j,c),\overrightarrow{t})\\ &\circledast\Phi(\eta(j,\lambda(a+b+c))\\ &\quad+\lambda\eta(j,a-2b+c) + \lambda\eta(j,a+b-2c) - 3\lambda\eta(j,a),\overrightarrow{t})\\ &\succeq \Phi(\alpha[\mu(j,\lambda(a+b+c)) - \lambda\mu(j,a) - \lambda\mu(j,b)\\ &\qquad-\lambda\mu(j,c)],\overrightarrow{t})\\ &\quad\circledast\Phi(\beta[\eta(j,\lambda(a+b+c)) + \lambda\eta(j,a-2b+c)\\ &\qquad+\lambda\eta(j,a+b-2c) - 3\lambda\eta(j,a)],\overrightarrow{t}),\end{aligned}$$

for any $\lambda \in \Xi$ and $\overrightarrow{t} \in (0,\infty)^n$. Since $\max\{|a|,|b|\} < 1$, then, $\eta(j,\lambda(a+b+c)) + \lambda\eta(j,a-2b+c) + \lambda\eta(j,a+b-2c) - 3\lambda\eta(j,a) = 0$ and $\mu(j,\lambda(a+b+c)) - \lambda\mu(j,a) - \lambda\mu(j,b) - \lambda\mu(j,c) = 0$ for any $\lambda \in \Xi$ and $\overrightarrow{t} \in (0,\infty)^n$. Thus, by Lemma 2, η and μ are $\mathbb{C}-$linear. According to (14) and the additive of μ and η,

$$\begin{aligned}&\Phi(\mu(j,[a,b,c]) - [\mu(j,a),b,c] - [a,\mu(j,b),c]\\ &\qquad-[a,b,\mu(j,c)],\overrightarrow{t})\\ &\circledast\Phi(\eta(j,[a,b,c]) - [\eta(j,a),\eta(j,b),\eta(j,c)],\overrightarrow{t})\\ &= \Phi(2^{3n}(\mu(j,\frac{[a,b,c]}{2^{3n}}) - [\mu(j,\frac{a}{2^n}),\frac{b}{2^n},\frac{c}{2^n}]\\ &-[\frac{a}{2^n},\mu(j,\frac{b}{2^n}),\frac{c}{2^n}] - [\frac{a}{2^n},\frac{b}{2^n},\mu(j,\frac{c}{2^n})]),\overrightarrow{t})\\ &\circledast\Phi(2^{3n}(\eta(j,\frac{[a,b,c]}{2^{3n}})\\ &\qquad-[\eta(j,\frac{a}{2^n}),\eta(j,\frac{b}{2^n}),\eta(j,\frac{c}{2^n})]),\overrightarrow{t})\\ &\succeq \phi(\frac{a}{2^n},\frac{b}{2^n},\frac{c}{2^n},\frac{\overrightarrow{t}}{2^n}),\end{aligned}$$

Which tend to $\mathbf{1}$ as $n \to \infty$. Hence, $\mu(j,[a,b,c]) - [\mu(j,a),b,c] - [a,\mu(j,b),c] - [a,b,\mu(j,c)] = 0$, and $\eta(j,[a,b,c]) - [\eta(j,a),\eta(j,b),\eta(j,c)] = 0$, for any $a,b,c \in \mho$. Therefore η is a FC-$\diamond$-ternary homomorphism and μ is a FC-$\diamond$-ternary derivation. ∎

Theorem 8. *Suppose $\phi : \mho^3 \circledast (0,\infty)^n \to diag\mathsf{H}_n(\mathfrak{C})$ and $\mu,\eta : \mho \to \mho$ are fuzzy operators satisfying* (13)*,* (14)*, $\mu(j,0) = \eta(j,0) = 0$ and*

$$\lim_{k\to\infty} \phi(2^k a, 2^k b, 2^k c, 2^k \overrightarrow{t}) = \mathbf{1}. \tag{15}$$

for any $\overrightarrow{t} \in (0,\infty)^n$. Then η is a FC-$\diamond$-ternary homomorphism and μ is a FC-$\diamond$-ternary derivation.

Proof: Through the same reasoning as in the proof of Theorem 7, μ and η are $\mathbb{C}-$linear. According to (14) and the additive of μ and η,

$$\begin{aligned}&\Phi(\mu(j,[a,b,c]) - [\mu(j,a),b,c] - [a,\mu(j,b),c]\\ &-[a,b,\mu(j,c)],\overrightarrow{t}) \circledast \Phi(\eta(j,[a,b,c])\\ &-[\eta(j,a),\eta(j,b),\eta(j,c)],\overrightarrow{t})\\ &= \Phi(\frac{1}{2^{3n}}(\mu(j,2^{3n}[a,b,c]) - [\mu(j,2^n a),2^n b,2^n c]\\ &-[2^n a,\mu(j,2^n b),2^n c]\\ &-[2^n a,2^n b,\mu(j,2^n c)]),\overrightarrow{t}) \circledast \Phi(\frac{1}{2^{3n}}(\eta(j,2^{3n}[a,b,c])\\ &-[\eta(j,2^n a),\eta(j,2^n b),\eta(j,2^n c)]),\overrightarrow{t})\\ &\succeq \phi(2^n a,2^n b,2^n c,2^n\overrightarrow{t}),\end{aligned}$$

which tends to $\mathbf{1}$, as $n \to \infty$. So $\mu(j,[a,b,c]) - [\mu(j,a),b,c] - [a,\mu(j,b),c] - [a,b,\mu(j,c)] = 0$, and $\eta(j,[a,b,c]) - [\eta(j,a),\eta(j,b),\eta(j,c)] = 0$, for any $a,b,c \in \mho$. Therefore η is a FC-$\diamond$-ternary homomorphism and μ is a FC-$\diamond$-ternary derivation. ∎

Theorem 9. *Suppose $\phi : \mho^3 \circledast (0,\infty)^n \to diag\mathsf{H}_n(\mathfrak{C})$ is a fuzzy control function such that there exists a $\varpi < 1$ with*

$$\lim_{k\to\infty} \phi(\frac{a}{2},\frac{b}{2},\frac{c}{2},\overrightarrow{t}) \succeq \phi(a,b,c,\frac{2\overrightarrow{t}}{\varpi}). \tag{16}$$

for any $\overrightarrow{t} \in (0,\infty)^n$. Suppose $\mu,\eta : \mho \to \mho$ are fuzzy operators satisfying (13)*,* (14)*, and $\mu(j,0) = \eta(j,0) = 0$. Then η is a FC-$\diamond$-ternary homomorphism and μ is a FC-$\diamond$-ternary derivation.*

Proof:

$$\lim_{n\to\infty} \phi(\frac{a}{2^n},\frac{b}{2^n},\frac{c}{2^n},\frac{\overrightarrow{t}}{2^n}) \succeq \phi(a,b,c,\frac{\overrightarrow{t}}{\frac{1}{2}\sum_{n=0}^{\infty}\varpi^n}) = \mathbf{1}.$$

∎

Theorem 10. *Suppose $\phi : \mho^3 \times (0,\infty)^n \to diag\mathsf{H}_n(\mathfrak{C})$ is a fuzzy control function such that there exists $a\varpi < 1$ with*

$$\lim_{k\to\infty} \phi(a,b,c,\overrightarrow{t}) \succeq \phi(\frac{a}{2},\frac{b}{2},\frac{c}{2},\frac{\overrightarrow{t}}{2\varpi}). \tag{17}$$

for any $\overrightarrow{t} \in (0,\infty)^n$. Suppose $\mu,\eta : \mho \to \mho$ are fuzzy operators satisfying (13)*,* (14)*, and $\mu(j,0) = \eta(j,0) = 0$. Then η is a FC-$\diamond$-ternary homomorphism and μ is a FC-$\diamond$-ternary derivation.*

Proof:

$$\lim_{n\to\infty} \phi(2^n a,2^n b,2^n c,2^n\overrightarrow{t}) \succeq \phi(a,b,c,\frac{\overrightarrow{t}}{\frac{1}{2}\sum_{n=0}^{\infty}\varpi^n}) = \mathbf{1}.$$

∎

Remark 3. *In* (14)*, the pair (η,μ) of a FC-$\diamond$-ternary homomorphism η and a FC-$\diamond$-ternary derivation μ can be replaced by the pair of a FC-$\diamond$-ternary homomorphism and a FC-$\diamond$-ternary homomorphism or the pair of a a FC-$\diamond$-ternary derivation and a FC-$\diamond$-ternary derivation.*

IV. Applications

Consider the one-parameter Mittag-Leffler function

$$\mathsf{M}_c(-\frac{\|\mathfrak{S}\|}{\mathfrak{Y}})=\sum_{k=0}^{\infty}\frac{(-\frac{\|\mathfrak{S}\|}{\mathfrak{Y}})^k}{\Gamma(1+ck)},$$

for any $c\in(0,1]$, $\mathfrak{S}\in\mathcal{W}$, and $\mathfrak{Y}\in(0,\infty)$, we have $(\mathcal{W},\mathsf{M}_c(-\frac{\|\mathfrak{S}\|}{\mathfrak{Y}}),\circledast_M)$ is a fuzzy normed space. For more details, you can see [1].

Now, consider the following increasing Hypergeometric function

$$\mathsf{H}(A,B,C;-\frac{\|\mathfrak{S}\|}{\mathfrak{Y}})=\sum_{k=0}^{\infty}\frac{(A)_k(B)_k}{(C)_k}\frac{(-\frac{\|\mathfrak{S}\|}{\mathfrak{Y}})^k}{k!},$$

in which $A,B,C\in[0,\infty]$, $\mathfrak{S}\in\mathcal{W}$, $\mathfrak{Y}\in(0,\infty)$.

We have $(\mathcal{W},\mathsf{H}(A,B,C;-\frac{\|\mathfrak{S}\|}{\mathfrak{Y}}),\circledast_M)$ is a fuzzy normed space.

Consider the following increasing Wright function

$$\mathsf{W}_{\imath,\jmath}(-\frac{\|\mathfrak{S}\|}{\mathfrak{Y}})=\sum_{k=0}^{\infty}\frac{(-\frac{\|\mathfrak{S}\|}{\mathfrak{Y}})^k}{k!\Gamma(\imath k+\jmath)},$$

in which $\imath,\jmath\in[0,\infty]$, $\mathfrak{S}\in\mathcal{W}$, $\mathfrak{Y}\in(0,\infty)$.

We have $(\mathcal{W},\mathsf{W}_{\imath,\jmath}(-\frac{\|\mathfrak{S}\|}{\mathfrak{Y}}),*_M)$ is a fuzzy normed space.

Now, we introduce the Fox's $\mathbb{H}$-function as follows

$$\mathbb{H}_{p,q}^{m,n}\left[z\middle|\begin{matrix}(a_j,A_j)_{1,p}\\(b_j,B_j)_{1,q}\end{matrix}\right]:=\frac{1}{2\pi i}\int_{\mathfrak{L}}\theta(\xi)z^{\xi}d\xi,\qquad(18)$$

$$i^2=1,\ z\in\mathbb{C}\backslash\{0\},\ z^{\xi}=\exp(\xi[\log|z|+i\ \arg(z)]),$$

in which $\log|z|$ denotes the natural logarithm of $|z|$ and $\arg(z)$ is not necessarily the principal value. For convenience,

$$\theta(\xi):=\frac{\prod_{j=1}^{m}\Gamma(b_j-B_j\xi)\prod_{j=1}^{n}\Gamma(1-a_j+A_j\xi)}{\prod_{j=m+1}^{q}\Gamma(1-b_j+B_j\xi)\prod_{j=n+1}^{p}\Gamma(a_j-A_j\xi)},$$

an empty product is interpreted as 1, and the integers m,n,p,q satisfy the inequalities

$$0\le n\le p\quad\text{and}\quad 1\le m\le q,$$

the coefficients

$$A_j>0\ (j=1,\dots,p)\quad\text{and}\quad B_j>0\ (j=1,\dots,q)$$

and the complex parameters

$$a_j\ (j=1,\dots,p)\quad\text{and}\quad b_j\ (j=1,\dots,q)$$

are so constrained that no poles of integrand in (18) coincide, and $\mathfrak{L}$ is a suitable contour of the Mellin-Barnes type (in the complex ξ-plane) which separates the poles of one product from those of the other.

we have $\mathbb{H}_{p,q}^{m,n}[-\frac{\|\mathfrak{S}\|}{\mathfrak{Y}}|_{(b_j,B_j)_{1,q}}^{(a_j,A_j)_{1,p}}]$ is a fuzzy Fox's $\mathbb{H}-$control function, for all $\mathfrak{S}\in\mathcal{W}$, $\mathfrak{Y}\in(0,\infty)$, $A_j,a_j>0$ $(j=1,\dots,p)$, $B_j,b_j>0$ $(j=1,\dots,q)$ and $p,q\in\mathbb{N}$.

Consider the following fuzzy control functions

$$\phi(a,b,\overrightarrow{t})=\operatorname{diag}\Big[\exp(-\frac{\mathsf{W}(\|a\|^{\mathsf{M}}+\|b\|^{\mathsf{M}})}{t}),\ \frac{t}{t+\mathsf{W}(\|a\|^{\mathsf{M}}+\|b\|^{\mathsf{M}})},\ \mathsf{M}_c(-\frac{\mathsf{W}(\|a\|^{\mathsf{M}}+\|b\|^{\mathsf{M}})}{t}),\ \mathbb{H}_{p,q}^{m,n}(-\frac{\mathsf{W}(\|a\|^{\mathsf{M}}+\|b\|^{\mathsf{M}})}{t}|_{(b_j,B_j)_{1,q}}^{(a_j,A_j)_{1,p}})\Big],$$

$$\phi(a,b,c,\overrightarrow{t})=\operatorname{diag}\Big[\mathsf{W}_{\imath,\jmath}(-\frac{\mathsf{W}(\|a\|^{\mathsf{M}}+\|b\|^{\mathsf{M}}+\|c\|^{\mathsf{M}})}{t}),\ \mathsf{H}(A,B,C;-\frac{\mathsf{W}(\|a\|^{\mathsf{M}}+\|b\|^{\mathsf{M}}+\|c\|^{\mathsf{M}})}{t})\Big],$$

where $\mathsf{M}>1$ and W is a nonnegative real number and $t\in(0,\infty)$.

They can replaced by the fuzzy control function ϕ presented in Theorem 1 and 2 and Theorem 7 and 8, respectively.

V. Conclusion

We introduced and solved two different additive-additive $(\alpha,\beta)-$Fuzzy functional inequalities. In addition, we investigated derivations and homomorphisms on matrix valued fuzzy Banach algebras and unital matrix valued FC-$\diamond$-algebras, associated to the above inequalities. Finally, we introduced new fuzzy control functions.

References

[1] S. R. Aderyani, and R, Saadati, Best approximations of the ϕ-Hadamard fractional Volterra integro-differential equation by matrix valued fuzzy control functions, Adv. Diff. Equations, 1(2021), pp. 1–21.

[2] P. Choonkil, X. Y. Wu, Homomorphism-derivation functional inequalities in C^*-algebras, AIMS Mathematics, 5(2020), 4482–4493.

[3] J. Bae, I. Chang, Some additive mappings on Banach C^*-algebras with derivation, J. Nonlinear Sci. Appl., 11 (2018), 335–341.

[4] Y. Ding, Ulam-Hyers stability of fractional impulsive differential equations, J. Nonlinear Sci. Appl., 11 (2018), 953–959.

[5] N. Eghbali, J. M. Rassias, M. Taheri, On the stability of a k-cubic functional equation in intuitionistic fuzzy n-normed spaces, Results Math., 70 (2016), 233–248.

[6] H. M. Srivastava, R. Saadati, and S. Y. Jang, Bi-additive σ–random operator inequalities and random quasi-$*$-multipliers on MB-algebras, Mathematical Sciences, 1(2021), pp. 1–12.

[7] C. Park, Additive C^*-functional inequalities and equations, J. Math. Inequal., 9 (2015), 17–26.

[8] C. Park, Additive C^*-functional inequalities in non-Archimedean normed spaces, J. Math. Inequal., 9 (2015), 397–407.

[9] Z. Wang, Stability of two types of cubic fuzzy set-valued functional equations, Results Math., 70 (2016), 1–14.

[10] Y. Jin, C. Park, and M. Th. Rassias, Hom-derivations in C^*-ternary algebras, Acta Math. Sin. (Engl. Ser.), 36 (2020), pp. 1025–1038.

[11] M. Madadi, D. O'Regan, Th. M. Rassias, and R. Saadati, Best approximation of κ-random operator inequalities in matrix MB-algebras, Journal of Inequalities and Applications, 1(2021), pp. 1–14.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Artificial Gorilla Troops Optimizer for Optimum Tuning of TID Based Power System Stabilizer

Mahmoud A. El-Dabah[1], Salah Kamel[2], Mohamed Khamies[2], Hossein Shahinzadeh[3], Gevork B. Gharehpetian[3]

[1] *Electrical Engineering Department, Al-Azhar University, Cairo, Egypt.*
[2] *Department of Electrical Engineering, Aswan University, Aswan 81542, Egypt.*
[3] *Department of Electrical Engineering, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran.*
dr_mdabah@azhar.edu.eg, skamel@aswu.edu.eg, mohamedahmedmak@yahoo.com, h.s.shahinzadeh@ieee.org, grptian@aut.ac.ir

***Abstract*— The artificial gorilla troops optimization (GTO) is a recently-introduced metaheuristic optimization technique for non-linear, non-convex and non-smooth problems with multiple types of variables. This algorithm, which is a nature-based technique, has been inspired by the social intelligence of gorilla troops. This study delves into how optimal can a power system stabilizer (PSS) unit be tuned incorporating GTO algorithm. The Integral Time Square Error (ITSE) was utilized as a fitness function, which desirably should be minimized. Four different controllers are employed to investigate the model of a single-machine scheme as a model for the infinite bus. The tilt-integral-derivative (TID) shows a better performance response in comparison with the PID controller, lead-lag controller, as well as fractional-order PID controller (FOPID) investigated. The results have implied that the GTO algorithm has obtained faster convergence over the other compared optimization techniques, while it has superiority in not trapping in local optimum points. It has shown also a high level of precision.**

***Index Terms* —SMIB, Tilt-integral-derivative (TID), Artificial gorilla troops optimization (GTO), power system stabilizer (PSS), Integral TimeSquare Error (ITSE).**

I. INTRODUCTION

Ensuring electrical power systems reliability and security calls for devoid of poorly damped oscillations. This condition is highly important in modern power grids especially with growing interconnections of power systems. Low-frequency oscillations can be a result of Weak interconnected large power systems when it is subjected even to small disturbance. These oscillations necessitate proper damping force to avoid power systems interruption. Moreover, these oscillations will affect synchronous machine lifetime besides the economic loss resulting from interruption. A power system stabilizer (PSS) is a well-known choice to remedy such impacts of these oscillations. The action of PSS is to initiate a rotor torque component in addition to remedying the phase lag error between the exciter input value and the electrical developed torque [1].

There are various types of PSS units such as Lead-lag controller, PI controller, PID controller, and fractional-order PID controller, known as FOPID. The lead-lag PSS is one of the conventional and widely used PSS due to its simplicity and easy implementation. This controller can tune the PSS unit using the linearized model of the electricity network in the vicinity of operating point. This method of tuning will not be the optimum parameters as per the dynamic topology and nature of the power grid operating point.

Metaheuristic optimization algorithms show noticed the wide application in different scale optimization problems. Various Metaheuristic algorithms are investigated in the optimum tuning of PSS taking into account a wide range of operational conditions like Jaya Algorithm [2], Grey Wolf Optimizer [3], Collective decision optimization algorithm [4], Bat Algorithm [5], kidney-inspired algorithm [6], sine cosine algorithm [7], Salp Swarm Algorithm [8], Cuckoo Search [9], whale optimization algorithm [10], Henry Gas Solubility Optimization Algorithm [11], hybrid algorithms like hybrid dynamic genetic algorithm (GA) and particle swarm optimization (PSO) [12], coyote optimization algorithm (COA) [13], Slime Mould Algorithm [14], farmland fertility algorithm [15].

The artificial gorilla troops optimization technique (GTO) is a recently proposed meta-heuristic optimization approach. This algorithm was inspired by the natural social behavior and social intelligence of the gorilla. According to the free lunch theorem, not all optimization algorithms do well in solving the optimization problem. Hence, the application of this algorithm

i.e., GTO for the optimum tuning of PSS will be evaluated in that study. The Integral Time Square Error (ITSE) will be utilized as a fitness function which is required to be minimized achieving the optimum tuning parameters of PSS. Moreover, a fractional-order calculus-based controller namely, Tilt-integral-derivative (TID) will be employed in PSS and compare its performance with PID, lead-lag controller, as well as FOPID to optimally tune a PSS. All these types of PSS will be engaged in the *Heffron Philips* model of a single synchronous machine to the infinite bus model (SMIB).

II. Problem Mathematical Modelling

The SMIB which is utilized in this paper includes a synchronous machine connected to a unified grid through a transmission line. A 4th order mathematical model of the synchronous machine is adopted. Fig.1 depicts SMIB with accompanied equipment like AVR and PSS. The following equations are used to represent the mathematical model of SMIB which is given in[16].

$$\dot{\delta} = \omega_b(\omega - 1) \tag{1}$$

$$\dot{\omega} = \frac{1}{M}(P_M - P_E - P_D) \tag{2}$$

$$\dot{E'_q} = \frac{1}{T'_{do}}\left[E_{fd} - E'_q - (x_d - x'_d)i_d\right] \tag{3}$$

$$\dot{E_{fd}} = \frac{1}{T_A}\left[K_A\left(v_{ref} - v_t + u_{PSS}\right) - E_{fd}\right] \tag{4}$$

$$P_E = E'_q i_q + \left(x_q - x'_d\right) i_d i_q \tag{5}$$

The linearized model of the SMIB can be recognized by the following state-space equations. Fig.2 depict these equations in a block diagram representation with the well known $K_1, K_2 \ldots K_6$ constants of Heffron Philips model.

$$\begin{bmatrix} \Delta\dot{\delta} \\ \Delta\dot{\omega} \\ \Delta\dot{E'_q} \\ \Delta\dot{E}_{fd} \\ \Delta\dot{v}_\omega \\ \Delta\dot{v}_s \end{bmatrix} = \begin{bmatrix} 0 & \omega_b & 0 & 0 & 0 & 0 \\ -\frac{K1}{2H} & -\frac{D}{2H} & -\frac{K2}{2H} & 0 & 0 & 0 \\ \frac{K4}{T'_{do}} & 0 & -\frac{1}{T'_{do}K3} & \frac{1}{T'_{do}} & 0 & 0 \\ -\frac{k_A K5}{T_A} & 0 & -\frac{k_A K6}{T_A} & -\frac{1}{T_A} & 0 & \frac{k_A}{T_A} \\ -\frac{K_1 K_{PSS}}{2H} & 0 & -\frac{K_2 K_{PSS}}{2H} & 0 & -\frac{1}{T_W} & 0 \\ -\frac{T_1 K_1 K_{PSS}}{2H\,T_2} & 0 & -\frac{T_1 K_2 K_{PSS}}{2H\,T_2} & 0 & \frac{1}{T_2}(1-\frac{T_1}{T_W}) & -\frac{1}{T_2} \end{bmatrix} \begin{bmatrix} \Delta\delta \\ \Delta\omega \\ \Delta E'_q \\ \Delta E_{fd} \\ \Delta v_\omega \\ \Delta v_s \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \frac{1}{2H} & 0 \\ 0 & 0 \\ 0 & \frac{k_A}{T_A} \\ \frac{K_{PSS}}{2H} & 0 \\ \frac{T_1 K_1 K_{PSS}}{2H\,T_2} & 0 \end{bmatrix} \begin{bmatrix} \Delta T_m \\ \Delta V_{ref} \end{bmatrix} \tag{6}$$

$$\begin{bmatrix} \Delta\omega \\ \Delta E_{fd} \\ \Delta V_t \\ \Delta T_e \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ K3 & 0 & K6 & 0 & 0 & 0 \\ K1 & 0 & K2 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta\delta \\ \Delta\omega \\ \Delta E'_q \\ \Delta E_{fd} \\ \Delta v_\omega \\ \Delta v_s \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta T_m \\ \Delta E_{fd} \end{bmatrix} \tag{7}$$

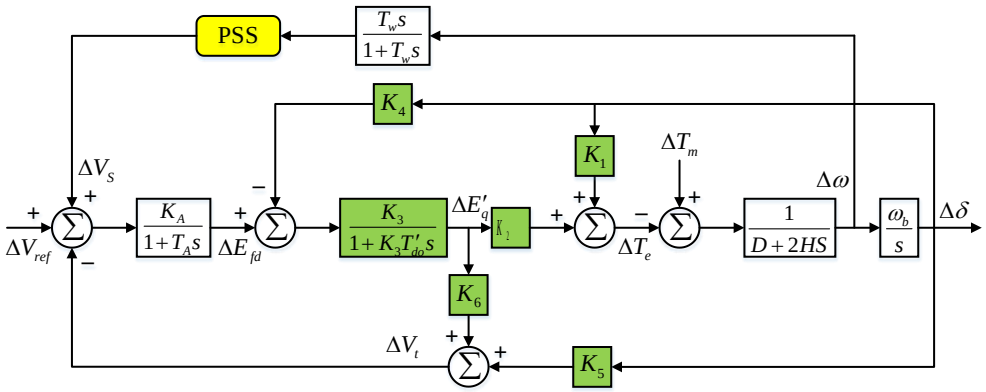


Figure 2. Heffron Phillips model of SMIB accompanied with PSS

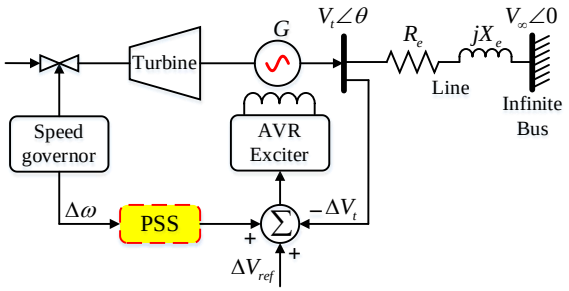


Figure 1. Schematic representation of SMIB with PSS

Based on the effective role of PSS which enhances the system oscillations damping of the electromechanical mode through the stabilization signal in the excitation system. Various types of PSS can be used to achieve this goal. In this research, the TID-based PSS will be exploited and compared with the PID and FOPID controllers. In Eq.(8), the transfer function of the TID controller is introduced while its schematic representation is shown in Fig.3.

$$TID = \frac{K_t}{s^{(\frac{1}{n})}} + \frac{K_I}{s} + K_D s \tag{8}$$

To examine the performance of these controllers and seeking for optimum parameters tuning of it, the ITSE is used as a fitness function to be minimized as given in Eq.(9).

$$ITSE = \int_0^{T_s} t \times e(t)^2 dt \tag{9}$$

The next section will spotlight the proposed investigated metaheuristic optimization algorithm to minimize the value of cost function in order to obtain the optimum parameters of PSS unit.

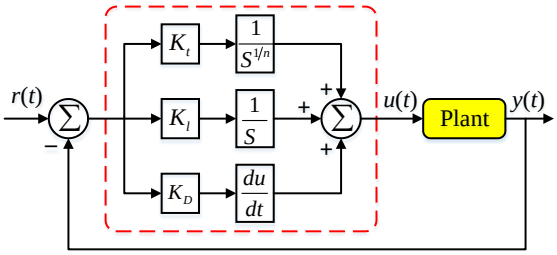


Figure 3. structure of TID controller

III. Artificial Gorilla Troops Optimizer: An Overview

The artificial Gorilla Troops Optimization (GTO) algorithm is a novel metaheuristic optimization technique [17]. The natural social intelligence of gorilla troops' is the main inspiration base for this algorithm. Five tactics describe the gorilla's natural behavior. Three of them are utilized in the exploration phase of the GTO optimizer. These strategies are migration to unidentified places, migration to identified places, and moving to the other gorilla. All these strategies are mathematically modeled by the following equations.

$$GX(it+1) = \begin{cases} (UB-LB)\times r_1 + LB & rand < p \\ (r_2 - C)\times X_r(it) + L \times H & rand \geq 0.5 \\ X(k) - L \times (L \times (X(it) - GX_r(it)) & \\ \quad + r_3 \times (X(it) - GX_r(it))) & rand < 0.5 \end{cases} \tag{10}$$

$$C = F \times \left(1 - \frac{it}{\text{Max}\_it}\right) \tag{11}$$

$$F = \cos(2 \times r_4) + 1 \tag{12}$$

$$L = C \times l \tag{13}$$

$$H = Z \times X(it) \tag{14}$$

$$Z = [-C, C] \tag{15}$$

where, UB, LB represents the variables' border limits. The selected and current position vector of the gorilla is denoted in iteration (it) by $GX(it+1)$ and $X(it)$ respectively while Max_it determines the maximum iteration number. Selecting the migration option probability is expressed by the parameter p which lie in the range [0,1]. At the end of the exploration stage, the solution $GX(it)$ will replace the $X(it)$ and turn out to be the silverback if fitness function value of $GX(it)$ is lower than the fitness function of $X(it)$.

On the other hand, competition for adult females and following the silverback constitute the tactics that are utilized in the exploitation phase of the GTO optimization algorithm. The mathematical formulation of these strategies can be given in Eqs. (16) to (22). Figure (4) presents the flow chart which depicts steps of execution of the GTO optimizer.

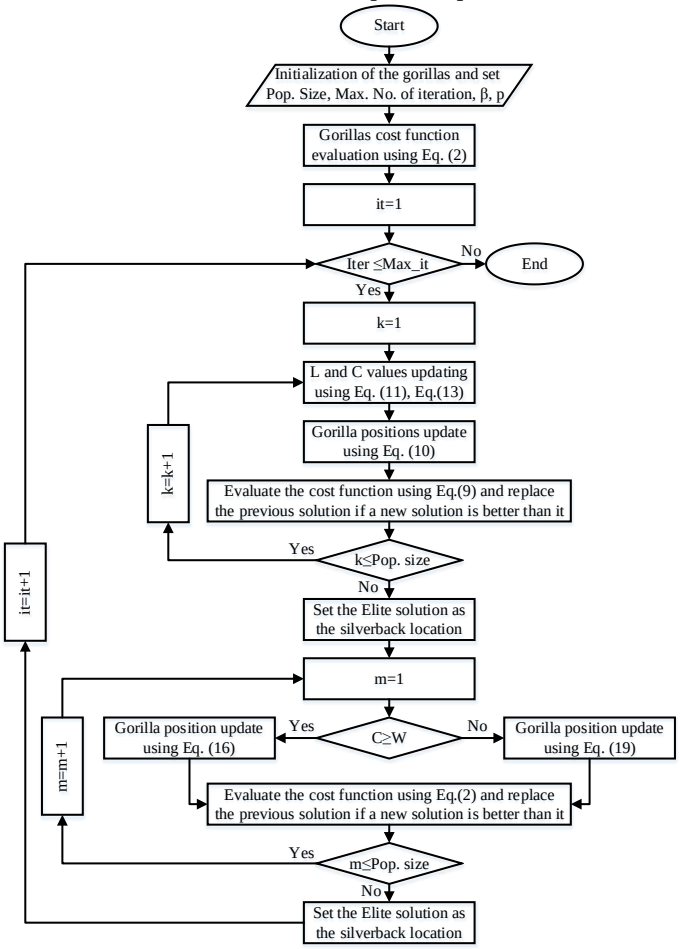


Figure 4. Flow chart of GTO

$$GX(it+1) = L \times M \times (X(it) - X_{silverback}) + X(it) \quad (16)$$

$$M = \left(\left| \frac{1}{N} \sum_{i=1}^{N} GX_i(it) \right| \right)^{\frac{1}{g}} \quad (17)$$

$$g = 2^L \quad (18)$$

$$GX(k) = X_{silverback} - (X_{silverback} \times Q - X(it) \times Q) \times A \quad (19)$$

$$Q = 2 \times r_5 - 1 \quad (20)$$

$$A = \beta \times E \quad (21)$$

$$E = \begin{cases} N_1 & rand \geq 0.5 \\ N_2 & rand < 0.5 \end{cases} \quad (22)$$

IV. Simulation Results and Discussion

Performance investigation of the proposed GTO optimization approach to optimally achieve the tunned parameters of PSS devices will be illustrated in this section. Moreover, the obtained results of TID-based PSS will be compared with the PID controller, lead-lag controller, as well as FOPID based PSS. Simulation results are employed with the MATLAB 2020a platform in an Intel core TM i7- 4790 CPU, 8 GB RAM Laptop. The parameters of the utilized SMIB model can be attained from [18]. Investigation of the system response upon the variation of 0.1 p.u in the mechanical torque T_m.

The TID-based PSS shows up the lowest ITSE value in comparison with PID controlling scheme, lead-lag controller, and FOPID based-PSS which is depicted in Fig.5 and Fig.6. Also, the change in the angular frequency in response to the simulated disturbance (i.e., change in T_m) was presented in Fig.7. The TID PSS performance indices indicate its superiority over the competitive PSSs as per the minimum settling time, MUS, and MOS which can be assured from Table.1 while Table.2 lists the optimum parameters of all examined structures of PSS.

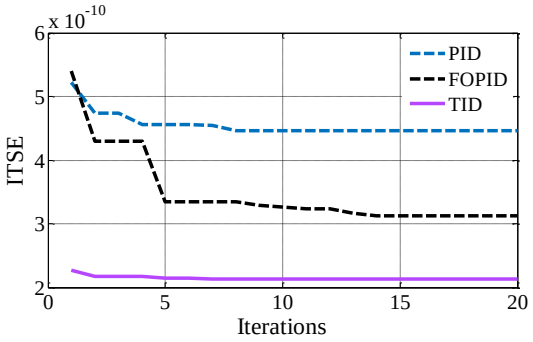


Figure 5. Convergence curves of investigated PSS controllers.

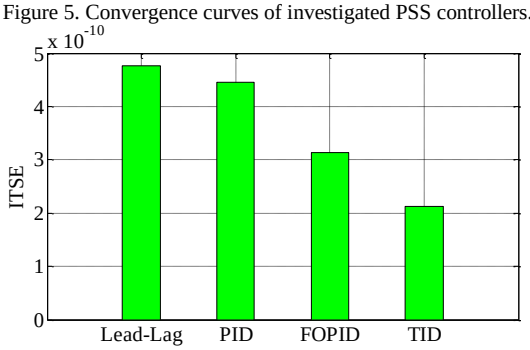


Figure 6. Bar plot of the ITSE value for investigated controllers

V. Conclusion

In this research, the artificial gorilla optimizer was adopted for the optimum tuning of the power system stabilizer. The Integral Time Square Error was used as a fitness function to be minimized seeking the optimum tuning parameters of the utilized PSSs. On the other hand, the lead-lag, PID, FOPID, and TID-based PSS controllers' performance were examined. The TID PSS shows superior performance regarding the lowest value for maximum overshoot and lowest value for settling time compared with the competitive PSS structures.

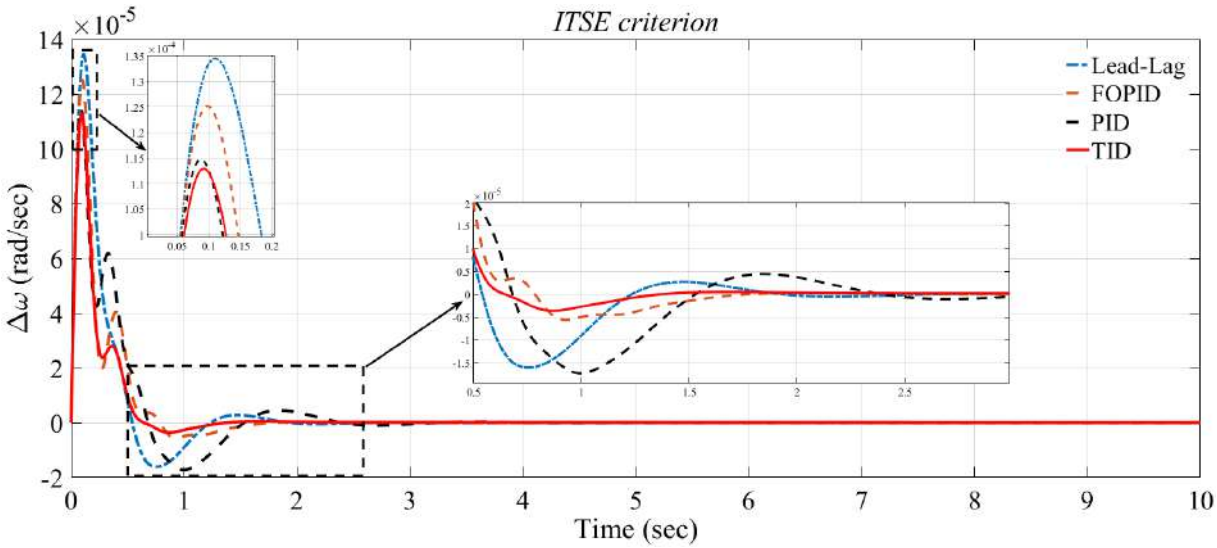


Figure 7. Angular frequency response corresponding to utilized controllers

TABLE I. Optimum parameters of utilized controllers using GTO.

	MOS	MUS	Settling time (sec)
Lead-lag	0.0001345	-1.608e-5	2.572
PID	0.0001148	-1.727e-5	2.969
FOPID	0.0001251	-5.566e-6	1.811
TID	0.0001128	-3.599e-6	1.335

TABLE II. PERFORMANCE INDICES OF UTILIZED CONTROLLERS USING GTO.

Controller	PSS controller parameters				
Lead-lag	$K_{PSS} = 15$		$T_1 = 0.548009$		$T_2 = 0.00100007$
PID	$K_P = 15$		$K_i = 15$		$K_d = 8.68463$
FOPID	$K_P = 15$	$K_i = 14.994$	$K_d = 15$	$\lambda = 0.1006$	$\mu = 0.8523$
TID	$K_t = 50$	$K_i = 50$	$K_d = 11.08$	$n = 50$	

REFERENCES

[1] Sauer, P. W., Pai, M. A., & Chow, J. H. (2017). *Power system dynamics and stability: with synchrophasor measurement and power system toolbox*. John Wiley & Sons.

[2] Sabo, A., Wahab, N. I. A., Othman, M. L., Jaffar, M. Z. A. M., Acikgoz, H., Nafisi, H., & Shahinzadeh, H. (2021). Artificial Intelligence-Based Power System Stabilizers for Frequency Stability Enhancement in Multi-machine Power Systems. *IEEE Access*.

[3] Shayeghi, H., Shayanfar, H. A., Asefi, S., & Younesi, A. (2016). Optimal Tuning and Comparison of Different Power System Stabilizers Using Different Performance Indices Via Jaya Algorithm. In *Proceedings of the International Conference on Scientific Computing (CSC)* (p. 34). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).

[4] Dey, P., Bhattacharya, A., & Das, P. (2017). Tuning of power system stabilizer for small signal stability improvement of interconnected power system. *Applied Computing and Informatics*. vol. 16, no. 1–2, pp. 3–28, doi: 10.1016/j.aci.2017.12.004.

[5] Sambariya, D. K., & Prasad, R. (2016). Design of optimal proportional integral derivative based power system stabilizer using bat algorithm. *Applied Computational Intelligence and Soft Computing*, *2016*.

[6] Khawaja, A. W., Kamari, N. A. M., & Zainuri, M. A. A. M. (2021). Design of a Damping Controller Using the SCA Optimization Technique for the Improvement of Small Signal Stability of a Single Machine Connected to an Infinite Bus System. *Energies*, *14*(11), 2996.

[7] Ekinci, S., & Hekimoglu, B. (2018, May). Parameter optimization of power system stabilizer via salp swarm algorithm. In *2018 5th international conference on electrical and electronic engineering (ICEEE)* (pp. 143-147). IEEE.

[8] Ekinci, S., Demiroren, A., & Hekimoglu, B. (2019). Parameter optimization of power system stabilizers via kidney-inspired algorithm. *Transactions of the Institute of Measurement and Control*, *41*(5), 1405-1417.

[9] Dasu, B., Sivakumar, M., & Srinivasarao, R. (2019). Interconnected multi-machine power system stabilizer design using whale optimization algorithm. *Protection and Control of Modern Power Systems*, *4*(1), 1-11.

[10] Chitara, D., Niazi, K. R., Swarnkar, A., & Gupta, N. (2018). Cuckoo search optimization algorithm for designing of a multimachine power system stabilizer. *IEEE Transactions on Industry Applications*, *54*(4), 3056-3065.

[11] Ekinci, S., İzci, D., & Hekimoğlu, B. (2021). Implementing the Henry Gas Solubility Optimization Algorithm for Optimal Power System Stabilizer Design.

[12] Guesmi, T., Alshammari, B. M., Almalaq, Y., Alateeq, A., & Alqunun, K. (2021). New Coordinated Tuning of SVC and PSSs in Multimachine Power System Using Coyote Optimization Algorithm. *Sustainability*, *13*(6), 3131.

[13] Abdulkhader, H. K., Jacob, J., & Mathew, A. T. (2018). Fractional-order lead-lag compensator-based multi-band power system stabiliser design using a hybrid dynamic GA-PSO algorithm. *IET Generation, Transmission & Distribution*, *12*(13), 3248-3260.

[14] Sabo, A., Abdul Wahab, N. I., Othman, M. L., Mohd Jaffar, M. Z. A., & Beiranvand, H. (2020). Optimal design of power system stabilizer for multimachine power system using farmland fertility algorithm. *International Transactions on Electrical Energy Systems*, *30*(12), e12657.

[15] Ekinci, S., Izci, D., Zeynelgil, H. L., & Orenc, S. (2020, October). An application of slime mould algorithm for optimizing parameters of power system stabilizer. In *2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)* (pp. 1-5). IEEE.

[16] Shankar, R., & Kundur, P. (1994). Power system stability and control II. *New York, McGraw-Hill Books pp581*.

[17] Abdollahzadeh, B., Soleimanian Gharehchopogh, F., & Mirjalili, S. (2021). Artificial gorilla troops optimizer: A new nature-inspired metaheuristic algorithm for global optimization problems. *International Journal of Intelligent Systems*, *36*(10), 5887-5958.

[18] El-Dabah, M. A., Kamel, S., Abido, M. A. Y., & Khan, B. (2021). Optimal tuning of fractional-order proportional, integral, derivative and tilt-integral-derivative based power system stabilizers using Runge Kutta optimizer. *Engineering Reports*, e12492.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)
March 2-4, 2022, Higher Education Complex of Bam

Assignment problem on fuzzy graphs

1st Javad Tayyebi, 2nd Hamid Bigdeli

1rd Department of Industrial Engineering, *Birjand University of Technology*, Birjand, I.R. Iran,
Javadtayyebi@birjandut.ac.ir
2st Institute for the Study of War, *Army Command and Staff University*, Tehran, I.R. Iran,
Hamidbigdeli92@gmail.com

***Abstract*— The assignment problem is the earliest combinatorial optimization problem, in which the decision maker has to assign some agents to some tasks in a way that the total cost is minimized. From the vision of graph theory, the problem is the same matching problem defined on a bipartite graph. This paper addresses a natural extension of the problem, in which the underlying graph is a fuzzy graph. This extension is meaningful in some applicable situations that there are one expert's opinions about the relative importance of any agent (task) and the relative satisfactory of any imputation. The goal of the decision maker is to make a decision not only regrading the costs, but also the expert's opinions. This paper concentrates on solving this problem. It first models the problem as a triobjective optimization. Then, the ϵ-constraint method is applied to propose an efficient algorithm for finding a compromise solution.**

***Index Terms*— Assignment problem, fuzzy graphs, satisfactory solution**

I. Introduction

The assignment problem contains a number m of agents as well as a number n of tasks. Any agent can be assigned to do any task, incurring some cost. Throughout this paper, it is supposed that $m \geq n$. However, the case $m < n$ can be analyzed in a similar fashion. By this assumption, the aim of the problem is to perform all tasks by assigning at exactly one agent to each task and at most one task to each agent, in such a way that the total cost of the assignment is minimized.

One can simply formulate the problem as

$$\min \quad \sum_{i\in V_1}\sum_{j\in V_2} c_{ij}x_{ij} \tag{1a}$$

$$\text{s.t.} \quad \sum_{j\in V_2} x_{ij} \leq 1 \qquad \forall i \in V_1, \tag{1b}$$

$$\sum_{i\in V_1} x_{ij} = 1 \qquad \forall j \in V_2, \tag{1c}$$

$$x_{ij} \in \{0,1\} \qquad \forall i \in V_1, j \in V_2, \tag{1d}$$

in which V_1 and V_2 are respectively the set of agents and the set of tasks. So, $|V_1| = m \geq n = |V_2|$. Here, c_{ij} is the cost of assigning the jth task to the ith agent.

Although formulation (1) is a zero-one linear programming, one can replace constraint (1d) to the nonnegativity constraint $x_{ij} \geq 0$ due to the unimodularity property of the problem [1]. This causes that the problem can be regarded as a linear programming problem, and can be solved by the well-known Simplex method [2].

From the graph theory viewpoint, the problem is defined in a weighted bipartite graph, and it consists of finding a matching which the sum of its arc weights is minimized. Figure 1 depicts the position of assignment problems among optimization problems.

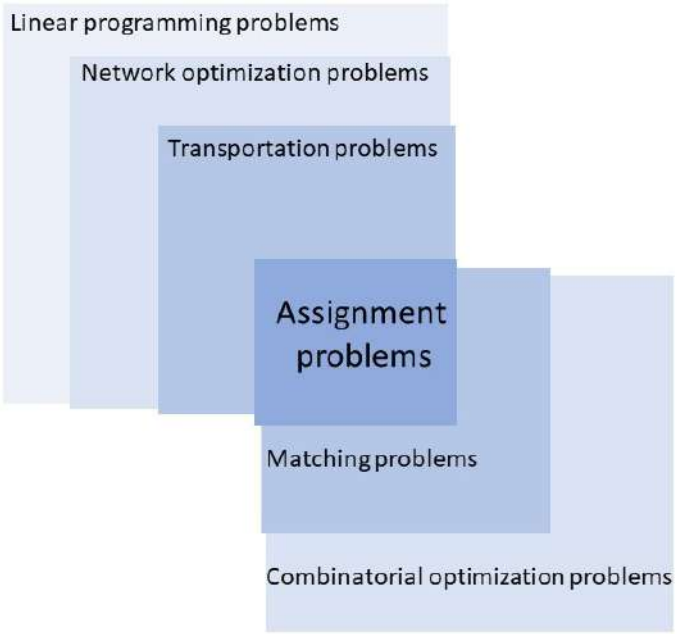


Fig. 1. The position of assignment problems among optimization problems

There are the plenty of papers which have studied problem (1) in fuzzy environments (see [3–6] for instance). All the papers have considered the problem in the case that costs c_{ij} are vague, and as a remedy, they are stated as fuzzy numbers. This paper is devoted to the different case that the underlying graph is a fuzzy graph. To the best our knowledge, this case is not studied until now in the literature. It has some real-world applications in the situations which there are an expert's opinions about the problem. The paper introduces the notion of solution, and presents an algorithm to obtain such a solution. The reminder of this paper is organized as follows. Section II provides some initial notions about fuzzy graphs. Section III presents the problem statement and furthermore, it develops the proposed algorithm to solve the problem. Section IV states some concluding remarks.

II. Preliminaries

This section states some initial notions and notations of fuzzy graphs used throughout the paper.

Let $G(V, A)$ be an directed graph in which V is the node set, and A is the arc set. Since this paper concentrates on assignment problems, it is assumed to G is bipartite, namely, the node set V can be partitioned into two set V_1 and V_2 so that $i \in V_1$ and $j \in V_2$ for every $(i, j) \in A$.

Let $X(\neq \emptyset)$ be a universal set. A fuzzy set (FS) $\tilde{A}$ is stated in term of a function from X to $[0, 1]$. It can be mathematically denoted as $\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) : x \in X\}$ in which $\mu_{\tilde{A}}(x)$ is the membership degree of x at $\tilde{A}$.

Suppose that $\tilde{V}$ is a fuzzy set on V and $\tilde{A}$ is also a fuzzy set on A. For simplicity notations, we denote by $\mu_{\tilde{A}}(i, j)$, instead of $\mu_{\tilde{A}}((i, j))$, the membership value of arc (i, j) in $\tilde{A}$. If

$$\mu_{\tilde{A}}(i, j) \leq \min\{\mu_{\tilde{V}}(i), \mu_{\tilde{V}}(j)\} \tag{2}$$

then, $\tilde{G} = (\tilde{V}, \tilde{A})$ is referred to as a fuzzy graph (FG), and $G(V, A)$ is called its underlying graph. In the special case that $\mu_{\tilde{V}}(i) = 0$ ($\mu_{\tilde{A}}(i, j) = 0$), we can remove node i (arc (i, j)) from the underlying graph G.

III. Problem statement

This section focuses on assignment problems in the case that the corresponding graph is a fuzzy one.

Assume $\tilde{G}(\tilde{V}, \tilde{A})$ is a bipartite fuzzy graph in which $\tilde{V} = \tilde{V}_1 \cup \tilde{V}_2$. The underlying graph of $\tilde{G}$ is $G(V_1 \cup V_2, A)$ in which V_1 is the set of agents and V_2 is the set of tasks. Here, A is the set of all allowable imputations, not necessarily all elements of $V_1 \times V_2$. $\mu_{\tilde{V}}(i)$ ($\mu_{\tilde{V}}(j)$) shows the relative importance of agent i (task j) from the viewpoint of an expert. Moreover, $\mu_{\tilde{V}}(i, j)$ is the relative satisfactory degree of the expert if agent i is assigned to task j. The inequality (2) states that the satisfactory degree of imputation (i, j) cannot be greater than the importance values of agent i and task j. The purpose of an assignment F is a maximal subset of A so that any two elements of F are not adjacent together.

Like the classic assignment problems, any imputation (i, j) is associated to a nonnegative crisp cost c_{ij}. As mentioned, we assume that the number of agents is greater than or equal to that of tasks, namely $m \geq n$. This assumption states that

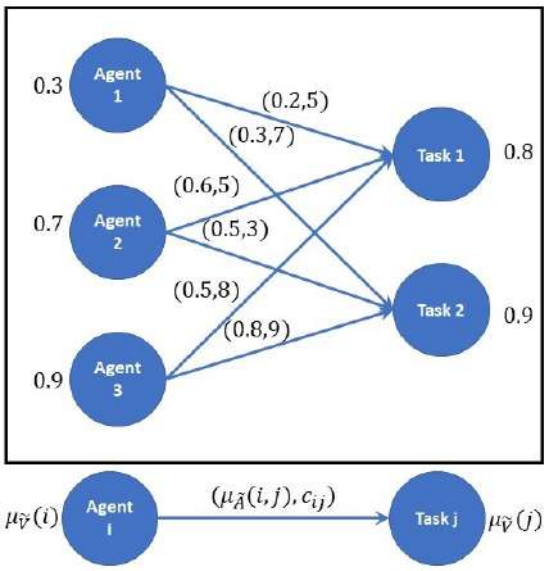


Fig. 2. An instance of the assignment problem defined on a fuzzy graph

all tasks are assigned while some agents may not be assigned. For this reason, we define an importance degree $sd(F)$ for any possible assignment F as

$$sd(F) = \min_{i:\exists(i,j)\in F}\{\mu_{\tilde{V}}(i)\}.$$

In the case that $n > m$, we can define $sd(F)$ in term of the membership values corresponding to tasks. Similarly, we define a satisfactory degree $id(F)$ for an assignment F as

$$id(F) = \min_{(i,j)\in F}\{\mu_{\tilde{A}}(i, j)\}.$$

By regarding the expert's opinions, the goal of solving the assignment problem is to find an assignment so that

- its total cost is minimized.
- its important degree is maximized;
- its satisfactory degree is maximized;

Example 1. *Figure 2 depicts an instance of the simple assignment problem containing 3 agents and 2 tasks. Like the classic version, if we only consider the costs, then* $F_1 = \{(Agent1, Task1), (Agent2, Task2)\}$ *is the optimal assignment with the total cost of 8. If the objective is to maximize the satisfactory degree, then* $F_2 = \{(Agent2, Task1), (Agent3, Task2)\}$ *is optimal while for the objective function of maximizing satisfactory degree, both the assignments* F_2 *and* $F_3 = \{(Agent2, Task2), (Agent3, Task1)\}$ *are optimal. Hence, we have to present an approach to present a compromise solution.*

Based on this argument, the problem can be formulated as

the following triobjective programming problem.

$$\max(-\sum_{i\in V_1}\sum_{j\in V_2} c_{ij}x_{ij}, id, sd), \tag{3a}$$

$$\sum_{j\in V_2} x_{ij} \le 1 \quad \forall i \in V_1, \tag{3b}$$

$$\sum_{i\in V_1} x_{ij} = 1 \quad \forall j \in V_2, \tag{3c}$$

$$id \le \mu_{\tilde{A}}(i,j) + M(1-x_{ij}) \qquad \forall (i,j) \in E, \tag{3d}$$

$$sd \le \mu_{\tilde{v}}(i) + M(1-\sum_{j\in V_2} x_{ij}) \qquad \forall i \in V_1, \tag{3e}$$

$$x_{ij} \in \{0,1\} \qquad \forall (i,j) \in E, \tag{3f}$$

To find a compromise solution of this problem, let us apply the ϵ-constraint method for the second and third objective functions. So, we have

$$z(sd, id) = \min \sum_{i\in V_1}\sum_{j\in V_2} c_{ij}x_{ij}, \tag{4a}$$

$$\sum_{j\in V_2} x_{ij} \le 1 \quad \forall i \in V_1, \tag{4b}$$

$$\sum_{i\in V_1} x_{ij} = 1 \quad \forall j \in V_2, \tag{4c}$$

$$id \le \mu_{\tilde{A}}(i,j) + M(1-x_{ij}) \qquad \forall (i,j) \in E, \tag{4d}$$

$$sd \le \mu_{\tilde{v}}(i) + M(1-\sum_{j\in V_2} x_{ij}) \qquad \forall i \in V_1, \tag{4e}$$

$$x_{ij} \in \{0,1\} \qquad \forall (i,j) \in E, \tag{4f}$$

Problem (4) is a parametric optimization problem with respect to the values id and sd. To solve this problem, we introduce an auxiliary graph $G_{sd,id}(V, A_{sd,id})$ for every fixed values $sd, id \in [0,1]$, in which the node set is the same set V, moreover, $A_{sd,id}$ contains an arc (i,j) if $\mu_{\tilde{V}}(i) \ge sd$ and $\mu_{\tilde{A}}(i,j) \ge id$. Based on this definition, the following result is immediate.

Lemma 1. *Any optimal assignment of $G_{sd,id}(V, A_{sd,id})$ is an optimal solution of problem* (4)*.*

Proof. The proof is straightforward.

Now, consider the optimal value of problem (4), that is $z(sd, id)$ as a bivariate function at $[0,1]\times[0,1]$. It is easy to see that it is a non-increasing step (piecewise constant) function with respect to both its variables. Let us focus more on this function. This function is not defined necessarily throughout $[0,1]\times[0,1]$. For example, $G_{sd,id}(V, A_{sd,id})$ does not contain any possible assignment when id tends to 1^-. So, the problem (4) is infeasible, and consequently, $z(sd, id)$ is undefined at this point. Moreover, the function is a trivial constant function when id tends to 0^+ because the optimal assignment of $G_{sd,id}(V, A_{sd,id})$ and $G(V, A)$ are the same. Note that this case yields only a weakly efficient solution of problem (3), and not an efficient solution. Based on these observations, we introduce an subinterval $[id_{min}, id_{max}] \subseteq [0,1]$ not containing these cases for the parameter id. id_{max} is the maximum value $id \in [0,1]$ so that there is at least one assignment in $G_{sd,id_{max}}(V, A_{sd,id_{max}})$ for some value $sd \in [0,1]$. id_{min} is the minimum membership value of arcs belonging to F where F is the optimal assignment with respect to the first objective function in $G(V, A)$.

A similar argument is valid for parameter sd. In this case, we introduce a desired subinterval $[sd_{min}, sd_{max}] \subseteq [0,1]$ for parameter sd. sd_{max} is the maximum value $sd \in [0,1]$ so that there is at least one assignment in $G_{sd_{max},id}(V, A_{sd_{max},id})$ for some values id. id_{min} is the minimum membership value of nodes belonging to F where F is the optimal assignment with respect to the first objective function in $G(V, A)$.

Based on the above discussion, hereafter, we limit the domain of $z(sd, id)$ to $[sd_{min}, sd_{max}] \times [id_{min}, id_{max}]$. To determine the behaviour of the function, it suffices to compute its values at its breakpoints. Let

$$(id_{min} =) id_1 < id_2 < \ldots < id_p (= id_{max})$$

be the breakpoints with respect to id and

$$(sd_{min} =) sd_1 < sd_2 < \ldots < sd_q (= sd_{max})$$

the breakpoints with respect to sd. The range of $z(id, sd)$ is the set

$$Range(z) = \bigcup_{i=1}^{q}\bigcup_{j=1}^{p}\{z(sd_i, id_j)\}.$$

The concept of our proposed method is to compute a compromise objective value z^* by the well-known mean value theorem for integrals. This can be obtained as

$$\begin{aligned} z^* &= \frac{1}{(sd_{max}-sd_{min})(id_{max}-id_{min})} \int_{sd_{min}}^{sd_{max}}\int_{id_{min}}^{idmax} z(sd, id)\mathrm{d}(id)\mathrm{d}(sd) \\ &= \frac{1}{(sd_{max}-sd_{min})(id_{max}-id_{min})} \sum_{i=2}^{p}\sum_{j=2}^{q} z(sd_{i-1}, id_{j-1})(sd_i - sd_{i-1})(id_i - id_{i-1}). \end{aligned} \tag{5}$$

Although z^* is a good candidate for the total cost, it is possible that this value does not correspond any function value $z(sd_i, id_j)$. Therefore, we choose the closest value z_{comp} of $Range(z)$ to z^*. So, the assignment corresponding to z_{comp} can be considered as the desired solution.

Now, we are ready to state our proposed algorithm for solving problem (3), formally.

IV. Conclusion

This paper considered the assignment problem defined on fuzzy graphs. The membership values of nodes were considered as the importance degree of agents/tasks from the viewpoint of an expert. Moreover, the membership values of arcs were considered as the satisfactory degree of the expert from the corresponding imputation.

At first, the problem was formulated as a tri-objective optimization problem. Then, using the ϵ-constraint method, the problem was converted into an ordinary assignment on an auxiliary graph. The objective value of the problem was seen as a bivariate function. Finally, a compromise solution was computed by the mean value theorem for integrals.

Algorithm 1 An algorithm to obtain a compromise assignment of problem (3)

Input: A bipartite fuzzy graph $\tilde{G} = (\tilde{V}_1 \cup \tilde{V}_2, \tilde{A})$ with arc costs c_{ij}.
Output: A compromise assignment F_{comp}.
Let M_V be the membership values of nodes at V_1 which are sorted in an increasing order.
Let M_A be the membership values of arcs which are sorted in an increasing order.
Use a binary search to find sd_{max} and id_{max} in M_V and M_A, respectively.
Using the well-known Hungarian algorithm, find an optimal assignment F on the underlying graph $G(V, A)$.
Let sd_{min} and id_{min} be respectively the minimum membership value of nodes and arcs belonging to F.
Update $M_V = M_V \cap [sd_{min} sd_{max}]$ and $M_A = M_A \cap [id_{min} id_{max}]$.
for $i = 1$ to $|M_V|$ **do**
    for $j = 1$ to $|M_A|$ **do**
        Compute $z_{ij} = z(sd_i, id_j)$ by Hungarian algorithm.
        Let F_{ij} be the assignment corresponding to z_{ij}.
    end for
end for
Compute z^* using (5).
Let $(i^*, j^*) = argmin_{i,j}\{|z^* - z_{ij}|\}$.
Set $F_{comp} = F_{i^* j^*}$.

As future works, it is meaningful to investigate the problem on the extensions of fuzzy graphs, such as the intuitive fuzzy graphs, and hesitant fuzzy graphs.

REFERENCES

[1] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, Network Flows: Theory, Applications and Algorithms, Prentice-Hall, Englewood Cliffs, New Jersey, USA Arrow, KJ, 1993.

[2] M. S. Bazaraa, and J. J. Jarvis, Linear programming and network flows, John Wiley and Sons, 1977.

[3] A. Emrouznejad, L. M. Zerafat Angiz, and W. Ho, "An alternative formulation for the fuzzy assignment problem", Journal of the Operational Research Society, vol. 63, pp. 59-63, 2012. +

[4] A. N. Gani, and V. N. Mohamed, "Solution of a fuzzy assignment problem by using a new ranking method", Intern. J. Fuzzy Mathematical Archive, vol. 2, pp. 8-16, 2013.

[5] P. S. Kumar, and R. J. Hussain, "A method for solving balanced intuitionistic fuzzy assignment problem", International Journal of Engineering Research and Applications, vol. 4, pp. 897-903, 2014.

[6] C. J. Lin, and U. P. Wen, "A labeling algorithm for the fuzzy assignment problem", Fuzzy sets and systems, vol. 142, pp. 373-391, 2004.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Best Proximity Point For Various Classes of Proximal Contraction Mapping In Fuzzy Metric Space

1st R.Gholipour, 2nd H.Mazaheri

1st Department of Mathematics, Yazd University, Yazd, Iran, gholipour.razieh@gmail.com.
2nd Department of Mathematics, Yazd University, Yazd, Iran, H.mazaheri@yazd.ac.ir

***Abstract*— We discuss the class of proximal contraction mappings, and we show that this singularity of best proximity point results for such mapping, is a directly concluded from it's fixed point theory.**

***Index Terms*— Best proximity point, Fixed point, Non-Archimedean fuzzy metric space.**

I. Introduction

The theory of fuzzy sets was offered by Zadeh [8] in 1965. Kramosil et al. [7] stated the consept of a fuzzy metric space in 1975. From 1975 up to now the notion of fixed point theorems and fuzzy metric space were studied, and proved by many authors. In 2013 C.Vetro et al. in [5] considered the problem of finding a least distance betwixt two nonempty sets in a fuzzy metric space. Recently, N.Hussain et al. [1] introduced various classes of proximal contraction mappings and obtained best proximity point theorems for suchlike fuzzy mapping in a non-Archimedean fuzzy metric space.

II. Preliminaries

At first, we briefly recall sevral basic notations and known results which will be need in the sequal discussion.

Definition 1. *Suppose $(X, M, \star)$ be a fuzzy metric space and let L, K be nonempty closed subsets of X. $L_0(s)$ and $K_0(s)$ Defined as following:*

$$L_0(s) := \{\lambda \in L : M(\lambda, \mu, s) = M(L, K, s),\ \textit{for some } \mu \in K\},$$

$$K_0(s) := \{\mu \in K : M(\lambda, \mu, s) = M(L, K, s).\ \textit{for some } \lambda \in L\},$$

wherever $M(L, K, s) = \sup\{M(\lambda, \mu, s) : \lambda \in L, \mu \in K\}$.

Definition 2 ([3])**.** *Suppose $T : X \to X$ and $\alpha : X \times X \to [0, +\infty)$ be two functions. If for all $a_1, b_1 \in X$,*

$$\alpha(a_1, b_1) \geq 1 \Rightarrow \alpha(Ta_1, Tb_1) \geq 1.$$

T is said to be α-admissible mapping.

Definition 3 ([6])**.** *Suppose $(X, M, \star)$ be a non-Archimedean fuzzy metric space and $\alpha : X \times X \to [0, +\infty)$ and $T : X \to X$ be an α-admisible mapping. Moreover, consider continuous functions $\kappa, \Phi : [0, 1] \to [0, 1]$ such that κ is decreasing function, $\kappa(s) > \kappa(1) - \Phi(1)$ and $\Phi(s) > 0$ for every $s \in (0, 1)$.If for any $a, b \in X$ and all $s > 0$,*

$$\alpha(a, Ta)\alpha(b, Tb)\kappa(M(Ta, Tb, s)) \leq \kappa(M(a, b, s)) - \Phi(M(a, b, s)).$$

T is called fuzzy $(\alpha - \beta - \kappa)$- contractive mapping

Theorem 1 ([2])**.** *Suppose $\rho : (0, +\infty) \to (0, 1)$ is a function and $\varphi \in \Phi$ and Suppose T be a self-mapping on $(X, M, \star)$. Inadition, assume that for every $a, b \in X$ via $a \neq b$ and any $s > 0$*

$$\varphi(M(Ta, Tb, s)) \leq \rho(s)\varphi(M(a, b, s)).$$

Then T has an only fixed point.

Inadition, $\varphi : [0, 1] \to [0, 1]$ is the set of all functions satisfying:
(φ_2) $\varphi(s) = 0$ if and only if $s = 1$.
(φ_1) φ is continuous and decreasing
Indicate by ϕ.

Theorem 2 ([5])**.** *Suppose L and K be a nonempty subsets of $(X, M, \star)$ such that $L_0(s)$ is nonempty and closed. Suppose non-self mapping $S : L \to K$ satisfies the following conditions:*
(i) $S(L_0(s)) \subseteq K_0(s)$
(ii) There is an only element $\varphi \in \Phi$ such that for every $\lambda, \mu, u, v \in L$ and every $s > 0$,

$$\left.\begin{array}{ll} M(u, S\lambda, s) & = M(L, K, s) \\ M(v, S\mu, s) & = M(L, K, s) \end{array}\right\} \Rightarrow \varphi(M(u, v, s)) \leq \rho(s)\varphi(M(\lambda, \mu, s)) \quad (1)$$

So there is an only element $a^\star \in L$ such that $M(a^\star, Sa^\star, s) = M(L, K, s)$ for all $s > 0$.

Theorem 3 ([6]). *Suppose $(X, M, \star)$ be a complete non-Archimedean fuzzy metric space, $\alpha : X \times X \to [0, +\infty)$, κ, Φ as in describtion 3 and T be a fuzzy $(\alpha - \beta - \kappa)$- contractive mapping so that the appendix claims is established:*
(i) $\exists a_0 \in X$ such that $\alpha(a_0, Ta_0) \geq 1$;
(ii) If $\{a_n\}$ is a sequence such that $\alpha(a_n, a_{n+1}) \geq 1$ for every $n \in N$, and $a_n \to a$ as $n \to +\infty$, then $\alpha(a, Ta) \geq 1$.
Therefore T has an only fixed point. Further, if $b = Tb$ points out $\alpha(b, Tb) \geq 1$ then T has an only fixed point.

Theorem 4 ([5]). *Suppose L and K be a nonempty subsets X such that $L_0(s)$ is nonempty. Suppose, non-self mapping$S : L \to K$ satisfying the conditions:*
(i) $S(L_0(s)) \subseteq K_0(s)$
(ii) There is a continuous function $\Phi : [0, 1] \to [0, 1]$, via $\Phi(s) > 0$ for every $s \in (0, 1)$, such that for every $a, b, u, v \in L$ and every $s > 0$ we have

$$\left.\begin{array}{ll} M(u, Sa, s) & = M(L, K, s) \\ M(v, Sb, s) & = M(L, K, s) \end{array}\right\}$$
$$\Rightarrow (M(u, v, s)) \geq M(a, b, s) + \Phi(M(a, b, s)) \qquad (2)$$

(iii) For every sequence $\{b_n\} \in K_0(s)$ and $a \in L$ satisfying $M(a, b_n, s) \to M(L, K, s)$ as $n \to +\infty$, there is an only element $k^\star \in L$ such that $M(k^\star, Sk^\star, s) = M(L, K, s)$ for every $s > 0$.

Definition 4 ([1]). *Suppose, $S : X \to X$, $\alpha^\star : X \times X \times (0, +\infty) \to [0, +\infty)$. If*
[1] for every $\beta, \gamma \in X$ and $s > 0$, $\alpha^\star(\beta, \gamma, s) \geq 1 \Rightarrow \alpha^\star(S\beta, S\gamma, s) \geq 1$,
[2] for every $\beta, \gamma, \nu \in X$ and $s > 0$, $\alpha^\star(\beta, \gamma, s) \geq 1, \alpha^\star(\gamma, \nu, s) \geq 1 \Rightarrow \alpha^\star(\beta, \nu, s) \geq 1$. The mapping S is called triangular $\alpha^\star$- admissible

Definition 5 ([1]). *Suppose, $S : L \to K$, $\alpha^\star : L \times L \times (0, +\infty) \to [0, +\infty)$. If the sequal conditions is hold, the mapping S is called triangular $\alpha^\star$-proximal admissible*
[1] S is $\alpha^\star-$ proximal admissible, that is, for every $\lambda_1, \lambda_2, \varrho_1, \varrho_2 \in L$ and $s > 0$,

$$\begin{cases} \alpha^\star(\lambda_1, \lambda_2, s) \geq 1 \\ M(\varrho_1, S\lambda_1, s) = M(A, B, s) \\ M(\varrho_2, S\lambda_2, s) = M(A, B, s) \end{cases} \Rightarrow \alpha^\star(\varrho_1, \varrho_2, s) \geq 1; \quad (3)$$

[2] for every $\beta, \gamma, \nu \in X$ and $s > 0$, $\alpha^\star(\beta, \gamma, s) \geq 1, \alpha^\star(\gamma, \nu, s) \geq 1 \Rightarrow \alpha^\star(\beta, \nu, s) \geq 1$

Definition 6 ([1]). *Suppose $S : L \to K, \alpha^\star : L \times L \times (0, +\infty) \to [0, +\infty)$.*
(a) If for every $\lambda_1, \mu_1, \varrho_1, \omega_1 \in L$ and $s > 0$,

$$\begin{cases} \alpha^\star(\lambda_1, \mu_1, s) \geq 1 \\ M(\varrho_1, S\lambda_1, s) = M(L, K, s) \\ M(\omega_1, S\mu_1, s) = M(L, K, s) \end{cases}$$
$$\Rightarrow M(\varrho_1, \omega_1, s) \geq M(\lambda_1, \mu_1, s) + \Phi(M(\lambda_1, S\mu_1, s)). \quad (4)$$

Therefore S is called modified $\alpha^\star - \Phi-$ proximal contractive mapping wherever Φ as in describtion 3.

(b) If for every $\lambda_1, \mu_1 \in X$ and $s > 0$, $\alpha^\star(\lambda_1, \mu_1, s) \geq 1$ implies

$$M(S\lambda_1, S\mu_1, s) \geq M(\lambda_1, \mu_1, s) + \Phi(M(\lambda_1, \mu_1, s)),$$

Therefore S is called modified $\alpha^\star - \Phi-$ contractive mapping

Theorem 5 ([1]). *Suppose $(X, M, \star)$ be a complete non-Archimedean fuzzy metric space and $S : X \to X$ satisfies this conditions:*
(a) S is a triangular $\alpha^\star$-admissible mapping;
(b) S is a modified $\alpha^\star - \Phi-$ contractive mapping;
(c) $\exists \lambda_0 \in X$ such that $\alpha^\star(\lambda_0, S\lambda_0, s) \geq 1$ for all $t > 0$;
(d) if $\{\lambda_n\}$ is a sequence in X such that $\alpha^\star(\lambda_n, \lambda_{n+1}, s) \geq 1, \forall n \geq 1$ and $\lambda_n \to \lambda$ as $n \to +\infty$, then $\alpha^\star(\lambda_n, \lambda, s) \geq 1, \forall n \geq 1$ and $s > 0$, Then S has a fixed point in X. Furthermore, if $\lambda = S\lambda$ and $\mu = S\mu$ to imply $\alpha^\star(\lambda, \mu, s) \geq 1, \forall t \geq 0$, then S has an only fixed point in X.

Theorem 6 ([1]). *Suppose $L, K \in C(X)$ such that $(X, M, \star)$ such that $L_0(s) \neq \Phi$ for all $s > 0$. Suppose, $S : L \to K$ satisfies in these conditions:*
(a) S is a triangular $\alpha^\star$-proximal admissible mapping and $S(L_0(s)) \subseteq K_0(s)$ for every $s > 0$;
(b) S is a modified $\alpha^\star - \Phi-$ proximal contractive mapping;
(c) if $\{\mu_n\}$ is a sequence in $K_0(s)$ and $\lambda_1 \in A$ are such that $M(\lambda_1, \mu_n, s) \to M(L, K, s)$, as $n \to +\infty$, then $\lambda_1 \in L_0(s)$ for every $s > 0$;
(d) $\exists$ $\lambda_0, \lambda_1 \in L_0(s)$ such that $M(\lambda_1, S\lambda_0, s) = M(L, K, s)$, $\alpha^\star(\lambda_0, \lambda_1, s) \geq 1$ for every $s > 0$
(v) if $\{\lambda_n\}$ is a sequence in X such that $\alpha^\star(\lambda_n, \lambda_{n+1}, s) \geq 1$ for every $n \geq 1$ and $\lambda_n \to \lambda$ as $n \to +\infty$, then $\alpha^\star(\lambda_n, \lambda, s) \geq 1$ for every $n \geq 1$ and $s > 0$. So $\exists \lambda^\star \in L$ so that $M(\lambda^\star, S\lambda^\star, s) = M(L, K, s)$, $\forall t > 0$. Furthermore, if $M(\lambda, S\lambda, s) = M(L, K, s)$ and $M(\mu, S\mu, s) = M(L, K, s)$ to imply $\alpha^\star(\lambda, \mu, s) \geq 1$, then S has an only best proximity point.

III. Application to fixed point theory

Now, we state the main conclusion.

Theorem 7. *Theorem 2 is directly concluded of Theorem 1.*

Proof: Suppose $\lambda \in L_0(s)$. Since $S(L_0(s)) \subseteq K_0(s)$ for all $s > 0$ there is $u \in L_0(s)$ such that $M(u, S\lambda, s) = M(L, K, s)$. u is unique. Suppose the contrary, there is another point $v \in L_0(s)$ such that $u \neq v$, $0 < M(u, v, s) < 1$ for every $s > 0$ and $M(v, S\lambda, s) = M(L, K, s)$. Now by using the condition (ii) of Theorem 2, there is an only element $\varphi \in \Phi$ such that

$$\left.\begin{array}{ll} M(u, S\lambda, s) & = M(L, K, s) \\ M(v, S\lambda, s) & = M(L, K, s) \end{array}\right\}$$
$$\Rightarrow \varphi(M(u, v, s)) \leq \rho(s)\varphi(M(\lambda, \lambda, s)) \leq \rho(s)\varphi(1) = 0$$
$$\varphi(M(u, v, s)) = 0 \Rightarrow M(u, v, s) = 1.$$

which is inconsistency by a assumption. Thus $M(u, v, s) \geq 1$ for all $s > 0$. Therefore, $v = u$. So for every point $a \in L_0(s)$ there is an only point $u \in L_0(s)$ such that $M(u, S\lambda, s) = M(L, K, s)$. Put $u = T(\lambda)$. Thus $T : L_0(s) \to L_0(s)$ is a mapping satisfying $M(T\lambda, S\lambda, s) = M(L, K, s)$. For all

$a \in L_0(s)$. We now claim that the self mapping $T : L_0(s) \to L_0(s)$ satisfies the following condition. Indeed, for any $a, b \in L_0(s)$, $s > 0$ we have

$M(T\lambda, S\lambda, s) = M(L, K, s)$ and $M(T\mu, S\mu, s) = M(L, K, s)$ again, since the condition ii of Theorem 2, we have

$$\varphi(M(T\lambda, T\mu, s)) \leq \rho(s)\varphi M(\lambda, \mu, s).$$

Now by using Theorem 1 T has an only fixed point $\nu \in L_0(s)$. This mean, $\nu = T(\nu)$ and therefore,

$$M(\nu, S\nu, s) = M(T\nu, S\nu, s) = M(L, K, s),$$

this mean S has an only best proximity point. ∎

Theorem 8. *Theorem 4 is directly concluded of Theorem 3.*

Proof: Assume $a \in L_0(s)$. Since $S(L_0(s)) \subseteq K_0(s)$ $\exists \eta \in L_0(s)$ for all $s > 0$ such that $M(\eta, Sa, s) = M(L, K, s)$. To see that η is unique assume there is another point $\zeta \in L_0(s)$ such that $\eta \neq \zeta$, $0 < M(\eta, \zeta, s) < 1$ any $s > 0$ and $M(\zeta, Sa, s) = M(L, K, s)$. Now by using assumption (b) of theorem 4, implies

$$\left.\begin{array}{ll} M(\eta, Sa, s) & = M(L, K, s) \\ M(\zeta, Sa, s) & = M(L, K, s) \end{array}\right\}$$
$$\Rightarrow M(\eta, \zeta, s) \geq M(a, a, s) + \Phi(M(a, a, s)) \qquad (5)$$

so $M(\eta, \zeta, s) = 1$.

That is, it contradicts the assumption, hence $M(\eta, \zeta, s) \geq 1$ for all $s > 0$, points that, $\zeta = \eta$. Then for any point $a \in L_0(s)$ there is an only point $\eta \in L_0(s)$ such that $M(\eta, Sa, s) = M(L, K, s)$. Set $\eta = T(a)$. So $T : L_0(s) \to L_0(s)$ is a mapping satisfying $M(Ta, Sa, s) = M(L, K, s)$, for all $a \in L_0(s)$. Then by taking $\alpha(a, b) = 1$ for all $a, b \in L_0(s)$ and $s > 0$ also $\psi(s) = 1 - s$ in the definition 3, and

$$\left.\begin{array}{ll} M(Ta, Sa, s) & = M(L, K, s) \\ M(Tb, Sb, s) & = M(L, K, s) \end{array}\right\}$$
$$\Rightarrow M(Ta, Tb, s) \geq M(a, b, s) + \Phi(M(a, b, s)), \qquad (6)$$

T is $(\alpha - \beta - \kappa)$-contractive mapping. It now follows from Theorem 3, that the mapping T has an only fixed point $\gamma \in L_0(s)$. That is, $\gamma = T(\gamma)$ and therefore,

$$M(\gamma, Sa, s) = M(T\gamma, Sa, t) = M(L, K, s),$$

which ensures that S has an only best proximity point. ∎

Theorem 9. *Theorem 6 is directly concluded of Theorem 5*

Proof: Suppose $\lambda \in L_0(s)$. Since $S(L_0)(s) \subseteq K_0(s)$, for all $s > 0$ there is $\mu \in L_0(s)$ such that $M(\mu, S\lambda, s) = M(L, K, s)$. We now claim that μ is unique. Suppose a contradiction $0 < M(\mu, \nu, s) < 1$ for all $s > 0$ and $\nu \neq \mu$ is elsewhere point of $L(s)$. Such that $M(\nu, S\lambda, t) = M(L, K, s)$ and $M(\mu, S\lambda, s) = M(L, K, s)$. Now from the condition (b) of Theorem 6, S is a modified $\alpha^\star - \Phi$-proximal contractive mapping. It follows that

$$M(\mu, \nu, s) \geq M(\lambda, \lambda, s) + \Phi(M(\lambda, \lambda, s)) \geq 1 + \Phi(M(\lambda, \lambda, s).$$

This implies $M(\mu, \nu, s) \geq 1$, which a contradiction. So $M(\mu, w, s) = 1$ for all $s > 0$. Hence, $\nu = \mu$. Then for every point $\lambda \in L_0(s)$ there is a unique point $\mu \in L_0(s)$ such that $M(\mu, S\lambda, s) = M(L, K, s)$. Put $\mu = T(\lambda)$. Therefore $T : L_0(s) \to L_0(s)$ is a mapping satisfying $M(T\lambda, S\lambda, s) = M(L, K, s)$. For all $\lambda \in L_0(s)$. We now claim that the self mapping $T : L_0(s) \to L_0(s)$ is a triangular $\alpha^\star$-admissible and modified $\alpha^\star - \Phi$-contractive mapping. By the fact that, for any $\lambda_1, \lambda_2 \in L_0(s)$, we have

$$\begin{cases} \alpha^\star(\lambda_1, \lambda_2, s) \geq 1 \\ M(T\lambda_1, S\lambda_1, s) = M(L, K, s), \\ M(T\lambda_2, S\lambda_2, s) = M(L, K, s), \end{cases} \Rightarrow \alpha^\star(T\lambda_1, T\lambda_2, s) \geq 1. \qquad (7)$$

We conclude that T is $\alpha^\star$-admissible. Also by using the fact that T is a modified $\alpha^\star - \Phi-$ contractive mapping, which implies that

$$M((T\lambda_1, T\lambda_2, s) \geq (M(\lambda_1, \lambda_2, s)) + \Phi(M(\lambda_1, \lambda_2, s)).$$

That is T is a modified $\alpha^\star - \Phi-$ contractive self-mapping. It now follows from Theorem 5, T has an only fixed point $\eta^\star \in L_0(s)$. This mean, $\eta^\star = T(\eta^\star)$ and Hence,

$$M(\eta^\star, S\lambda, s) = M(T\eta^\star, S\lambda, s) = M(L, K, s).$$

This yields S has an only one best proximity point. ∎

CONCLUSIONS

Existence, convergence and singularity of best proximity points for various classes of non-self mappings has attracted the attention mathematicians. It may be checked as extention of fixed point theorems. But, some existence conclutions may be obtained from the same fixed point theorems directly, and we can not consider such conclusions as real extention of fixed point theory.

REFERENCES

[1] Hussain, Nawab and Kutbi, MA and Salimi, Peyman, *Global optimal solutions for proximal fuzzy contractions*, *Physica A: Statistical Mechanics and its Applications*, **551**, 123925, 2020, Elsevier

[2] Shen, Yonghong and Qiu, Dong and Chen, Wei, *Fixed point theorems in fuzzy metric spaces*, *Applied Mathematics Letters*, **25**, **2**, 138–141, 2012, Elsevier

[3] Samet, Bessem and Vetro, Calogero and Vetro, Pasquale, Fixed point theorems for α–ψ-contractive type mappings, Nonlinear analysis: theory, methods & applications, 75, 4, 2154–2165, 2012, Elsevier

[4] George, A and Veeramani, P, *On some results in fuzzy metric spaces*, *Fuzzy sets and systems*, **64**, **3**, 395–399, 1994, Elsevier

[5] Vetro, Calogero and Salimi, Peyman, *Best proximity point results in non-Archimedean fuzzy metric spaces*, *Fuzzy Information and Engineering*, **5**, **4**, 417–429, 2013, Taylor & Francis

[6] Salimi, Peyman and Vetro, Calogero and Vetro, Pasquale *Some new fixed point results in non-Archimedean fuzzy metric spaces*, *Nonlinear Analysis: Modelling and Control*, **18**, **3**, 344–358, 2013

[7] Kramosil, Ivan and Michálek, Jiří *Fuzzy metrics and statistical metric spaces*, *Kybernetika*, 11, 5, 336–344, 1975
[8] Lotfi Zadeh, A, Fuzzy sets, Information and control, 8, 3, 338–353, 1965.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Communication-based Optimization Algorithm: A Meta-heuristic Technique for Solving Single-Objective Problems

Majid Hadi, Reza Ghazizadeh

Faculty of Electrical and Computer Engineering, *University of Birjand*, *Birjand, Iran*
majid.h71@birjand.ac.ir, rghazizade@birjand.ac.ir

***Abstract*— In this paper, a science-inspired meta-heuristic optimization algorithm is proposed and called communication-based optimization algorithm (COA). The COA originates from the power allocation policy to users in non-orthogonal multiple access (NOMA)-based wireless communication networks. Four main steps of COA such as channel gain of users, power allocation, velocity of users and movement mechanism are modeled for equipping the COA with high exploration and exploitation phases. The COA is evaluated with standard mathematical test problems and compared with PSO, GSA, MFO, SSA, and FDA. The results show that the proposed algorithm is very competitive compared to well-known algorithms and in some cases its performance outperforms other meta-heuristic techniques, significantly.**

***Index Terms*—Meta-heuristic, communication-based optimization algorithm, power allocation, NOMA, channel gain**

I. Introduction

In recent years, meta-heuristic optimization algorithms have been considered by researchers as primary methods to solve the complex engineering problems. The reasons for the popularity of these methods can be summarized in three main reasons [1, 2]: 1) Simplicity due to easy implementation and simple inspiration; 2) can be used for a wide range of optimization problems; 3) local optima avoidance due to the stochastic nature of meta-heuristic algorithms.

It is hard to find an algorithm that provides the best solution for all set of optimization problems. In other words, an algorithm may offer a promising solution for a set of problems, but it may shows poor performance for other sets. This encourages researchers to propose new optimization methods or modify existing algorithms. Based on this, many meta-heuristic algorithms have emerged in recent decades.

The population-based meta-heuristic algorithms can be divided into four groups: evolution-based, physics-based, swarm-based, and human-based. Evolution-based methods are inspired by biological evolution, such as reproduction, mutation, recombination, and selection. Search agents during the optimization process play the role of individuals in a population, and the fitness function specifies the quality of the candidate solutions. The most popular algorithm introduced in this domain is Genetic Algorithm (GA) [3], which simulates the Darwin's theory of evolution. Other popular algorithms include Genetic Programming (GP) [4], Biogeography-Based Optimizer (BBO) [5] and Gene Expression Programming (GEP) [6]. The second group of population-based algorithms uses the physical rules for mathematical modeling. Some of the algorithms in this field are: Central Force Optimization (CFO) [7], Artificial Physics Optimization (APO) [8], Gravitational Search Algorithm (GSA) [9], and Gravitational Interactions Optimization (GIO) [10].

The third group of algorithms imitates the behavior of animals in the nature. The most famous algorithm in this field is Particle Swarm Optimization (PSO) [11], which simulates the social behavior in flock of birds in migration. Other popular algorithms can be pointed to: Ant Colony Optimization (ACO) [12] mimicking the behavior of real ants when they search for food, Whale Optimization Algorithm (WOA) [1] following the behavior of humpback whales, Grey Wolf Optimizer (GWO) [2] which imitates the leadership hierarchy and hunting mechanism of grey wolves and Salp Swarm Algorithm (SSA) [13], which formulates the swarming behavior of salps when navigating and foraging in oceans.

Finally, the fourth group of population-based algorithms originates from the human behaviors and laws in the world. Some of the algorithms in this domain can be pointed to: Virus Optimization Algorithm (VOA) [14] simulating the behaviour of viruses attacking a living cell, Harmony Search (HS) [15] following musician's behaviors in improvisation process, League Championship Algorithm (LCA) [16] inspiring by sport championships, and Soccer League Competition (SLC) [17].

In this work, we propose a population-based algorithm

called Communication-based Optimization Algorithm (COA). The COA models the approach of power allocation to multiple users based on the non-orthogonal multiple access (NOMA) in wireless communication networks. Channel gain of users, power allocation policy, and movement mechanism of users are three main steps of the proposed algorithm. The COA algorithm is tested with 13 standard test functions. The results illustrate that the COA algorithm provide better performance compared to well-known algorithms in the most cases due to the integrating of exploration and exploitation mechanisms.

The rest of this paper is organized as follows: Section II describes the proposed algorithm. The results and discussion of test functions are presented in Sections III. Finally, Section IV concludes the paper.

II. Communication-based optimization algorithm (COA)

In this section, first we provide the introduction of cellular networks and NOMA protocol then present the mathematical model of the COA algorithm.

A. NOMA Protocol

In the cellular communication networks, geographical area is divided into many cells and each cell contains one base station (BS) so that to data transmission all users connect to the BS. In this communication network, to access the medium, different multiple access techniques are introduced.

One of the potential and promising multiple access techniques in the fifth-generation (5G) wireless networks is the non-orthogonal multiple access (NOMA) protocol which can provide the high spectral efficiency compared to the traditional techniques [18–20]. The main idea behind the NOMA is that one frequency band can be allocated to multiple users. Then, the successive interference cancellation (SIC) at the receiver side is implemented to interference mitigation, user detection and decoding [18–20]. For further explanation of the SIC technique, let us consider a general K-user downlink NOMA system, as shown in Fig. 1, where the BS uses the one frequency band to transmit the information signals to the all users at the same time. Thus, in this system, each user will receive its message and other messages of users as interference signals, simultaneously. In the receiver side, to extract the desired message, the SIC first appears the stronger interferences, and subtracts them from the received signal. Since each user receives the all transmitted messages (desired and interfering messages), the power allocation mechanism for efficient SIC in downlink NOMA system is essential. As the results presented in [18–20], the user with the higher channel gain (lower distance to the BS) should transmit with the lower power level whereas the message with lower user channel gain (higher distance to the BS) transmits with higher power level.

To model the mathematical communication-based inspiration, assume that G_i denotes the channel gain between user i and BS. Also, assume that the channel gain of K-user system is sorted as follows

$$G_1 > G_2 > ...G_i > ... > G_K. \tag{1}$$

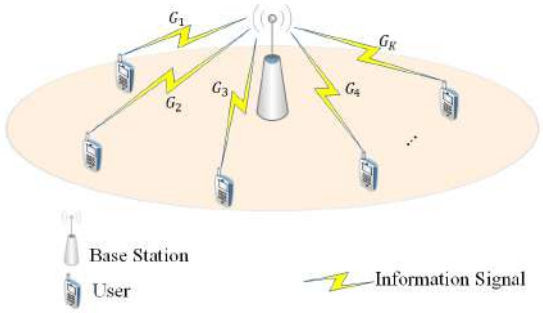


Fig. 1. Downlink NOMA system.

To perform efficient and successful SIC operation, the following condition must be met for the power allocation policy

$$P_1 < P_2 < ...P_i < ... < P_K. \tag{2}$$

where P_i is the transmitted power of BS to user i. It is observed that power allocation policy is proportion to the reverse of the channel gain of users.

In order to generalize this inspiration to optimization algorithms, cell, users and BS denote the search space, search agents and global solution, respectively. Users want to get closer to the BS to minimize the transmitted power of the BS. In this work, the NOMA communication-based feature is mathematically modeled.

B. Mathematical Model

In this sub-section, the mathematical model of channel gain of users, power allocation policy, and movement mechanism is first provided. The pseudo code of COA algorithm is then presented.

1) Channel gain of users: In wireless communication networks, the channel gain depends on the distance between the user and the BS. Since the position of the BS in the search space is unknown a priori, we assume that the user with the current best fitness as BS. After the BS is defined, the channel gain between other users and BS can be modeled as

$$G_i^t = rand \times d_{i,BS^t}^{-a}. \tag{3}$$

where G_i^t illustrates the channel gain between user (search agent) i and BS (current best fitness) at iteration t, $rand$ is a random value with uniform distribution in the interval [0,1] and expresses channel coefficient, d_{i,BS^t} denotes the distance between user i and BS, and a is a fixed value which expresses path loss exponent in the wireless communications. It should be noted that the BS position should be updated on every iteration if there is a better fitness.

2) Power allocation policy: As stated, to perform efficient SIC, the transmitted power should be assigned according to the reverse of the channel gain. Thus, we have the following equation

$$P_i^t = \frac{\beta}{G_i^t}. \tag{4}$$

where P_i^t indicates the power of the BS for data transmission to user i at iteration t and β is a constant value.

In practice, the channel gains are small values and increasing the interval range provides smaller channel gains. Therefore, the allocated powers to users will be large values. Based on the simulation experiences, these powers are not suitable for the proposed algorithm and create chaos in the search space. Hence, the normalized powers can provide an appropriate solution for the proposed COA algorithm. Thus, the COA algorithm uses the normalized powers as

$$P_i^t = \frac{P_i^t}{max(\overrightarrow{P^t})}. \tag{5}$$

where $\overrightarrow{P^t}$ is the power vector of users at iteration t, namely $\overrightarrow{P^t} = [P_1^t, P_2^t, \ldots, P_{popsize}^t]^T$. $popsize$ denotes the population size or number of users in the search space. Thus, the power allocation policy to all users can be modeled as follows

$$P_i^t = \begin{cases} P_{min} & \text{for BS} \\ min(max(P_{min}, P_i^t), P_{max}) & \text{for other users} \end{cases} \tag{6}$$

where P_{min} and P_{max} represent the minimum and maximum powers of each user, respectively which are predefined values. It is seen that BS has the minimum allocated power compared to other users because it presents the best solution in the search space or is close to optimum.

3) Movement mechanism: After the channel gain and power allocation strategies are formulated, the users must be move towards the BS (best solution) by using a mechanism and update their position. For this purpose, a physical formulation is applied. In physics, power is the rate of doing work or the amount of transferred or converted energy per time unit. Also, power, as a function of time, is the rate (i.e. derivative) at which work is done, thus can be expressed as follows

$$P_i^t = \frac{dW}{dt}. \tag{7}$$

where W and t denote the work and time, respectively. Because the time in optimization algorithms is defined in the form of iterations, we use the iteration term instead of the time. Furthermore, the work is defined as a force F applied over a distance x, namely $W = F.x$. Thus, for a constant F, the above equation can be rewritten as

$$P_i^t = F.\frac{dx_i^t}{dt} = F.V_i^t. \tag{8}$$

where F linearly increases from 0 to $1.4k_1$ over the iterations and $k_1 = 100/popsize$. V_i^t is the velocity of user$i$ at iteration t. Based on (8), the velocity of user i at iteration t can be calculated as $V_i^t = \frac{P_i^t}{F}$.

Now, based on equation $V = at + V_0$, where V_0 is the initial velocity, the acceleration of the user i at iteration t is given as follows

$$a_i^t = \frac{P_i^t}{F} - V_i^{t-1}. \tag{9}$$

where V_i^{t-1} is the velocity of the user i at iteration $t-1$. To update the position of the users, the following equation is utilized (Newton's law of motion)

$$x_i^t = \frac{1}{2}a_i^t t^2 + V_i^{t-1}t + x_i^{t-1}. \tag{10}$$

TABLE I
PSEUDO CODE OF THE COA ALGORITHM

```
Initialize the first population of users randomly
Calculate the fitness of users
Find the user with the best fitness and assume it is BS
for t = 1 to maximum number of iterations
  for each user
    Calculate distance between the each user and BS
    Calculate the channel gain between the each user and BS
    based on (3)
    Calculate the required power of users based on equation (6)
    Calculate the velocity of users based on (12)
    Update the position of users using (13)
  end for
  Compute the fitness of all users
  Update the BS if a user has the better fitness than the BS
end for
Return BS
```

where x_i^{t-1} specifies the position of user i at iteration $t-1$. By replacing and simplifying, the COA algorithm uses the following equation to update the position of users.

$$x_i^t = \frac{P_i^t}{F}(x_{BS}^t - x_i^{t-1}) + w.V_i^{t-1} + x_i^{t-1}. \tag{11}$$

where x_{BS}^t is the position of BS (best solution) at current iteration, w is linearly decreased from 1.2 to 0.5 over the course of iterations. Basically, the constants F and w can adjust the accuracy levels of exploitation and exploration phases, respectively. It should be noted that in the process of simplification, the factor of $\frac{1}{2}$ is neglected. According to (11), it can be seen that the velocity of the user i at iteration t is equal to

$$V_i^t = \frac{P_i^t}{F}(x_{BS}^t - x_i^{t-1}) + w.V_i^{t-1}. \tag{12}$$

and,

$$x_i^t = V_i^t + x_i^{t-1}. \tag{13}$$

where $x_{min} \leq x_i^t \leq x_{max}$. x_{min} and x_{max} are the lower and upper bounds of the search space, respectively.

C. Pseudo Code of COA

In the COA algorithm, the users are initialized randomly in the search space. The initial velocity of users are considered zero. The user with the best fitness is assumed as BS. In each iteration, channel gain, allocated power, velocity for each user is calculated based on (3), (6) and (12), respectively. Then, the position of each user is updated using (13). After computing the fitness of all users, the user with best fitness is introduced as the BS in the next iteration if its fitness is better than the fitness of the current BS. The optimization process ends when the iteration reaches the maximum number of iterations. The pseudo code of the COA algorithm are presented in Table I.

III. RESULTS AND DISCUSSION

In this section, the numerical results of the COA algorithm are evaluated by solving 13 standard test problems. These problems are classical benchmark functions used by many researchers [1, 2, 13]. The test problems are divided into

TABLE II
RESULTS OF UNIMODAL TEST FUNCTIONS

F	Metric	COA	PSO [11]	GSA [9]	MFO [21]	SSA [13]	FDA [22]
F1	ave	**0**	8.403	1.93e-17	1000.0	9.31e-9	2.31e-11
	std	**0**	6.343	5.05e-18	2999.9	2.51e-9	5.65e-11
F2	ave	**0**	1.978	2.22e-8	30.337	0.553	5.68e-9
	std	**0**	0.906	3.39e-9	22.574	0.599	4.14e-9
F3	ave	**0**	358.075	291.77	18637.1	43.557	0.566
	std	**0**	150.695	83.522	11418.1	38.358	0.258
F4	ave	**0**	7.448	0.018	51.997	3.856	14.811
	std	**0**	2.248	0.099	11.089	1.714	3.463
F5	ave	**28.739**	290.403	33.339	18373.7	270.990	29.077
	std	**0.029**	153.658	26.751	35838.5	516.155	23.496
F6	ave	**0**	43.200	**0**	1993.3	9.081e-4	3.71e-11
	std	**0**	12.573	**0**	3986.9	1.993e-9	1.17e-10
F7	ave	**0.00018**	0.0276	0.0199	2.216	0.0515	0.054
	std	**0.00026**	0.0153	0.0093	4.109	0.0156	0.021

TABLE III
RESULTS OF MULTIMODAL TEST FUNCTIONS

F	Metric	COA	PSO [11]	GSA [9]	MFO [21]	SSA [13]	FDA [22]
F8	ave	-7282.41	-6119.2	-2927.0	-8528.1	-7686.3	**-8552.634**
	std	849.36	870.63	**392.25**	792.32	687.07	719.375
F9	ave	**0**	48.818	15.057	144.32	49.416	53.429
	std	**0**	12.204	3.547	43.339	16.468	15.102
F10	ave	**2.66e-15**	3.645	3.67e-9	15.008	1.372	3.519
	std	**2.71e-15**	0.620	4.89e-10	7.537	0.969	0.759
F11	ave	**0**	1.036	3.719	9.053	0.007	0.040
	std	**0**	0.109	1.445	27.129	0.009	0.039
F12	ave	0.081	3.642	**0.052**	0.238	3.376	0.715
	std	**0.035**	1.511	0.063	0.473	2.069	0.976
F13	ave	0.596	4.873	**0.0003**	0.618	0.998	0.052
	std	0.533	4.038	**0.002**	0.783	4.607	0.121

TABLE IV
AVERAGE REQUIRED TIME TO SOLVE THE PROBLEM (IN SECOND)

COA	PSO [11]	GSA [9]	MFO [21]	SSA [13]	FDA [22]
0.7172	1.4970	5.5015	0.5648	0.6295	1.7124

two groups: unimodal and multimodal test problems, which are provided in Appendix. Unimodal test functions have an optimum point that evaluates the exploitation operation of an optimization algorithm. Multimodal test functions have more than one optimum point (global optimum and local optima), and thus they have more complexity than unimodal functions. Therefore, these functions verify the exploration operation and escape from the local optima of an algorithm. To confirm the results, the COA algorithm is compared to well-known algorithms such as PSO, GSA, Moth-Flame Optimization (MFO) [21], SSA and, Flow Direction Algorithm (FDA) [22]. For all the algorithms, 50 number search agents and 1000 iterations are used for solving the test function problems. Furthermore, in simulation setup, we set the minimum power of users P_{min} to 0.001, maximum power of users P_{max} to 0.65, parameter a to 0.5, and β to 3. The average value of the minimum costs (ave) and standard deviation (std) over 30 independent runs are reported in Tables II-III.

The test functions F1-F7 are unimodal, which evaluate the exploitation capability of an algorithm. Table II shows that the COA algorithm provides the optimal solution for 5 test functions, F1-F4 and F6 over all runs. For other test functions, namely F5 and F7, the proposed algorithm provides better solution compared to the other algorithms. Therefore, these results show that the COA can present marvelous exploitation compared to the studied algorithms. As a result, this algorithm is very effective for solving the convex optimization problems.

The results on multimodal test functions (F8-F13) are provided in Table III. The reported results show that the COA algorithm outperforms other studied algorithms on most multimodal functions. Also, it is observed that the proposed algorithm is able to obtain the optimal solution for 3 test functions, F9, F10 and F11 over all runs. In fact, these results indicate the COA has very good exploration due to integrated mechanisms discussed in the previous section.

In Table IV, COA is compared with other algorithms in terms of average consumed time to solve the unimodal and multimodal test problems. For all the algorithms, 30 design variables and 1000 iterations are considered. Also, the results are obtained by Core i5 CPU and 8.00 GB RAM. As can be observed from the data in the table, the proposed algorithm is very competitive compared to well-known algorithms and in some cases its performance outperforms other meta-heuristic techniques such as PSO, GSA and FDA.

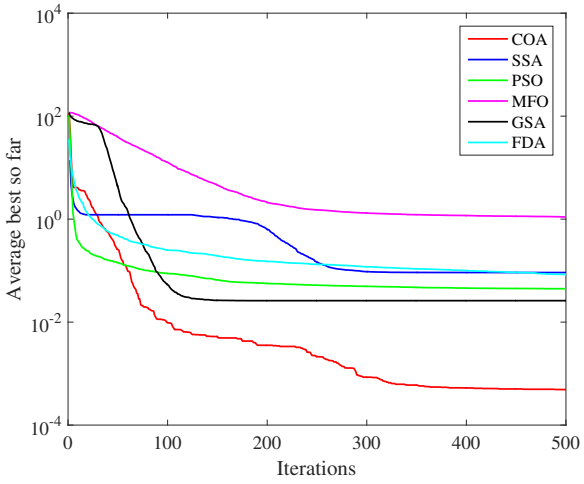


Fig. 2. Comparison of convergence of COA, SSA, PSO, MFO, GSA, and FDA for test function F7.

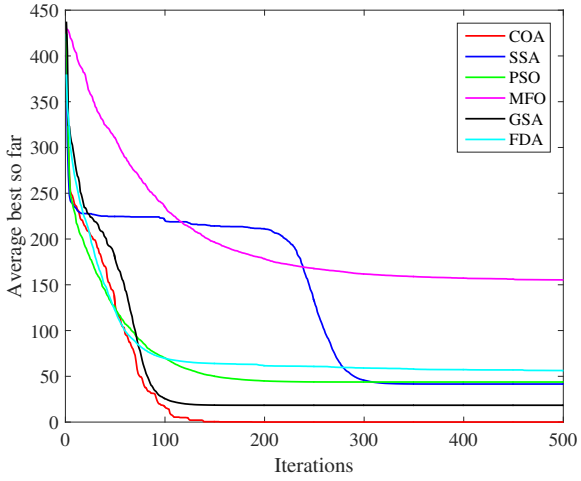


Fig. 3. Comparison of convergence of COA, SSA, PSO, MFO, GSA, and FDA for test function F9.

It is observed that the COA algorithm is able to guide the search agents (users) to promising regions of the search space and exploit the best one due to the efficient exploitation mechanism. Therefore, this issue guarantees convergence of the proposed algorithm. The convergence plots of the COA, SSA, PSO, MFO, GSA, and FDA are compared and provided in Fig. 2 and Fig. 3 for test functions F7 and F9, respectively. For all the algorithms, 50 search agents, 500 iterations and 30 design variables are utilized. Please note that average best so far illustrates the average best fitness so far at each iteration over 30 independent runs. As the curves in the figure show, the COA can explores promising areas of the search space to find the best solution and exploits it carefully due to the efficient exploration and exploitation operations.

IV. Conclusion

This paper proposed a meta-heuristic optimization algorithm called COA. The COA mimicked the power allocation strategy in NOMA-based wireless communication networks. Several operators such as channel gain, power allocation and movement policy were mathematically modeled for equipping the proposed algorithm with high exploitation and exploration operations. The performance of the COA was evaluated with 13 standard test functions. To confirm the results, the COA was compared to well-known algorithms such as PSO, GSA, MFO, SSA, and FDA. First, the results on unimodal test functions showed that the proposed algorithm has very good exploitation and can provide optimal solution for most of the test problems. Second, the obtained results on the multimodal test functions verified the exploration capability of the COA algorithm.

Appendix

The unimodal and multimodal test functions are provided in Table V.

In the Table V, y_i is defined as follows

$$y_i = 1 + \frac{1+x_i}{4}u(x_i, a, k, m) \tag{14}$$

TABLE V
Test Functions

Test function	Dim.	Range	f_{min}
$F_1(x) = \sum_{i=1}^{n} x_i^2$	30	[-100,100]	0
$F_2(x) = \sum_{i=1}^{n} \lvert x_i \rvert + \prod_{i=1}^{n} \lvert x_i \rvert$	30	[-10,10]	0
$F_3(x) = \sum_{i=1}^{n} \left(\sum_{j-1}^{i} x_j\right)^2$	30	[-100,100]	0
$F_4(x) = max_i \left(\lvert x_i \rvert, 1 \leq i \leq n\right)$	30	[-100,100]	0
$F_5(x) = \sum_{i=1}^{n-1} [100 \left(x_{i+1} + x_i^2\right)$ $+(x_i - 1)^2]$	30	[-30,30]	0
$F_6(x) = \sum_{i=1}^{n} \left([x_i + 0.5]\right)^2$	30	[-100,100]	0
$F_7(x) = \sum_{i=1}^{n} ix_i^4 + random[0,1)$	30	[-1.28,1.28]	0
$F_8(x) = \sum_{i=1}^{n} -x_i \sin\left(\sqrt{\lvert x_i \rvert}\right)$	30	[-500,500]	-418.98 29*Dim
$F_9(x) = \sum_{i=1}^{n} [x_i^2 - 10\cos(2\pi x_i)$ $+10]$	30	[-5.12,5.12]	0
$F_{10}(x) =$ $-20exp\left(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}\right)$ $-exp\left(\frac{1}{n}\sum_{i=1}^{n} \cos 2\pi x_i\right) + 20 + e$	30	[-32,32]	0
$F_{11}(x) = \frac{1}{4000}\sum_{i=1}^{n} x_i^2$ $+\prod_{i=1}^{n} \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	30	[-600,600]	0
$F_{12}(x) = \frac{\pi}{n}[10\sin(\pi y_1)$ $+\sum_{i=1}^{n-1} \left(y_i - 1\right)^2$ $[1 + 10\sin^2(\pi y_{i+1})] + (y_n - 1)^2]$ $+\sum_{i=1}^{n} u(x_i, 10, 100, 4)$	30	[-50,50]	0
$F_{13}(x) = 0.1[-\sin^2(3\pi x_1)$ $+\sum_{i=1}^{n} (x_i - 1)^2$ $[1 + \sin^2(3\pi x_i + 1)]$ $(x_n - 1)^2 \sin^2(2\pi x_n)]$ $+\sum_{i=1}^{n} u(x_i, 5, 100, 4)$	30	[-50,50]	0

References

[1] S. Mirjalili and A. Lewis, "The Whale Optimization Algorithm," Adv. Eng. Softw., vol. 95, pp. 51–67, 2016.

[2] S. Mirjalili, S. Mohammad, and A. Lewis, "Grey Wolf Optimizer," Adv. Eng. Softw., vol. 69, pp. 46–61, 2014.

[3] J. H. Holland, "Genetic Algorithms," Sci. Am., vol. 267, pp. 66–72, 1992.

[4] J. R. Koza, "Genetic programming," 1992.

[5] D. Simon, "Biogeography-Based Optimization," IEEE

Trans. Evol. Comput., vol. 12, pp. 702–713, 2008.
[6] C. Ferreira, “Gene Expression Programming: A New Adaptive Algorithm for Solving Problems,” Complex Syst., vol. 13, pp. 87–129, 2001.
[7] R. Formato, “Formato RA . Central force optimization: A new metaheuristic with applica- tions in applied electromagnetics,” Prog. Electromagn. Res., vol. 77, pp. 425–491, 2007.
[8] A. Gorbenko and V. Popov, “The Force Law Design of Artificial Physics Optimization for Robot Anticipation of Motion,” Adv. Stud. Theor. Phys., vol. 13, pp. 625–628, 2012.
[9] E. Rashedi, H. Nezamabadi-pour, and S. Saryazdi, “GSA: A Gravitational Search Algorithm,” Inf. Sci. (Ny)., vol. 179, pp. 2232–2248, 2009.
[10] J. J. Flores, R. L´opez, and J. Barrera, “Gravitational Interactions Optimization,” Learn. Intell. Optim., pp. 226–237, 2011.
[11] J. Kennedy and R. Eberhart, “Particle Swarm Optimization,” Proc. IEEE Int. Conf. Neural Networks, vol. 4, pp. 1942–1948, 1995.
[12] M. Dorigo, “Optimization, Learning and Natural Algorithms,” Ph.D. Thesis, Politec. di Milano, Italy, 1992.
[13] S. Mirjalili, A. H. Gandomi, S. Zahra, and S. Saremi, “Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems,” Adv. Eng. Softw., vol. 114, pp. 163–191, 2017.
[14] Z. W. Geem and J. H. Kim, “A New Heuristic Optimization Algorithm: Harmony Search,” Simulation, vol. 76, pp. 60–68, 2001.
[15] A. H. Kashan, “League Championship Algorithm: A New Algorithm for Numerical Function Optimization,” 2009 Int. Conf. Soft Comput. Pattern Recognit., pp. 43–48, 2009.
[16] N. Moosavian and B. K. Roodsari, “Soccer League Competition Algorithm , a New Method for Solving Systems of Nonlinear Equations,” Int. J. Intell. Sci., vol. 4, pp. 7–16, 2014.
[17] N. Moosavian and B. Kasaee, “Soccer league competition algorithm: A novel meta-heuristic algorithm for optimal design of water distribution networks,” Swarm Evol. Comput., vol. 17, pp. 14–24, 2014.
[18] M. Hadi and R. Ghazizadeh, “Joint sub-carrier allocation and 3D beamforming design in OMA-NOMA based mmWave heterogeneous networks under channel uncertainties,” AEU-International J. Electron. Commun., vol. 137, p. 153809, 2021.
[19] Z. J. Ali, N. K. Noordin, A. Sali, and F. Hashim, “Fair Energy-Efficient Resource Allocation for Downlink NOMA Heterogeneous Networks,” IEEE Access, vol. 8, pp. 200129–200145, 2020.
[20] A. Masaracchia, L. D. Nguyen, T. Q. Duong, C. Yin, O. A. Dobre, and E. Garcia-Palacios, “Energy-Efficient and Throughput Fair Resource Allocation for TS-NOMA UAV-Assisted Communications,” IEEE Trans. Commun., vol. 68, no. 11, pp. 7156–7169, 2020.
[21] S. Mirjalili, “Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm,” Knowledge-Based Syst., vol. 89, pp. 228–249, 2015.
[22] H. Karami, M. V. Anaraki, S. Farzin, and S. Mirjalili, “Flow Direction Algorithm (FDA): A Novel Optimization Approach for Solving Optimization Problems,” Comput. Ind. Eng., vol. 156, p. 107224, 2021.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Computing The Range Of Optimal Values Of Interval Linear Programming Problems: Comparing Genetic Algorithm With Monto Carlo Simulation

1st Moslem Javanmard, 2nd Hassan Mishmast Nehi

1st Math teacher with a PhD, *Education Department*, Zabol, Iran, javanmard.moslem@gmail.com
2nd Mathematics Department, *University of Sistan and Baluchestan*, Zahedan, Iran, hmnehi@hamoon.ac.ir

***Abstract*— In many real world problems, system parameters or model coefficients may be bounded between lower and upper bounds due to a variety of uncertainties. Over the past decades, intensive research efforts have focused on interval problems by two sub-models to tackle such uncertainties. In most of methods, interval problems by two sub-models (best and worst models) with deterministic parameters are formulated. Computing the best value is easy, but obtaining the worst value is very complicated or even impossible. In this article, we obtain the optimal range of interval problem via genetic algorithm and present one illustrative example to verify and compare the obtained results of it with the obtained results through Monte Carlo simulation and the exact optimal values.**

***Index Terms*— exact optimal values; genetic algorithm; interval linear programming; Monte Carlo simulation**

I. Introduction

Decision making in many real world problems in uncertain environments (such as: portfolio selection problem [7], inventory management [4]) are complicated. Therefore, in most of methods, these problems have been modeled with the help of a number of crisp or interval linear programming (ILP) problems [2, 11–16]. In some of them, ILP problems by two sub-models (best and worst models) with deterministic parameters are formulated [2, 3, 9, 12, 13]. The best and the worst method (BWC) proposed by Tong converts ILP model into two sub-models [9]. Also, genetic algorithm (GA) is recently developed optimization technique and rapidly become popular. So, many researchers, including Lin, et al [8], Wu, et al [10] considered GA to deal with such problems. GA is the generic computational technique espoused from the progression of biological life in the natural humanity [6]. Actually, GA is a global optimization algorithm by simulating heredity and process of evolution in environment. It uses three operating process that are selection, crossover and mutation to be survival of the fittest. Computing the best value is easy, but still, despite a lot of methods, obtaining the worst value is too complicated or even impossible. In this paper, to eliminate these complexities we obtain the optimal range of ILP problems with GA.

II. Overview of ILP problem

An interval number $x^{\pm}$ is generally shown with $[x^-, x^+]$ where $x^- \leq x^+$. If $x^- = x^+$, then $x^{\pm}$ will be degenerate. An interval vector is defined as $b^{\pm} = [b^-, b^+] = \{b^- \leq b \leq b^+\}$, where b^- and b^+ are two numbers in $\mathbb{R}$ and $b^- \leq b^+$. We denote the center and the radius of $b^{\pm}$ with $b^c = \frac{1}{2}(b^+ + b^-)$, and $\Delta_{b^{\pm}} = \frac{1}{2}(b^+ - b^-)$, respectively. Also, an interval matrix is defined as $A^{\pm} = [A^-, A^+] = \{A^- \leq A \leq A^+\}$, where A^- and A^+ are two matrices in $\mathbb{R}^{m\times n}$, and $A^- \leq A^+$. We denote the center and the radius of $A^{\pm}$ with $A^c = \frac{1}{2}(A^+ + A^-)$, and $\Delta_{A^{\pm}} = \frac{1}{2}(A^+ - A^-)$, respectively. The set of all $m \times n$ interval matrices is denoted by $\mathbb{IR}^{m\times n}$. A square matrix $A^{\pm}$ is called regular if each $A \in A^{\pm}$ is non-singular. An interval vector as $x^{\pm} = [x^-, x^+] = \{x^- \leq x \leq x^+\}$, where $x^-, x^+ \in \mathbb{R}^n$ is one column interval matrix.

A general form of ILP model with equality constraints can be expressed as follows:

$$\begin{aligned} &\text{Min } Z^{\pm} = c^{\pm}x^{\pm} \\ &\text{s.t.} \\ &\qquad A^{\pm}x^{\pm} = b^{\pm}, \\ &\qquad x^{\pm} \geq 0, \end{aligned} \tag{1}$$

where the right-hand sides of the constraints and the coefficients of both objective function and constraints are interval matrices (i.e., $b^{\pm}, c^{\pm}, A^{\pm}$). Or, equivalently:

$$\begin{aligned} &\text{Min } z^{\pm} = \sum_{j=1}^{n} c_j^{\pm} x_j^{\pm} \\ &\text{s.t.} \\ &\qquad \sum_{j=1}^{n} a_{ij}^{\pm} x_j^{\pm} = b_i^{\pm}, \quad i = 1, 2, \ldots, m, \\ &\qquad x_j^{\pm} \geq 0, \quad j = 1, 2, \cdots, n, \end{aligned} \tag{2}$$

where the coefficients are equal, $c_j^{\pm} = [c_j^-, c_j^+]$, $a_{ij}^{\pm} = [a_{ij}^-, a_{ij}^+]$, $b_i^{\pm} = [b_i^-, b_i^+]$ and $z^{\pm} = [z^-, z^+]$.

The interval $[z_{opt}^-, z_{opt}^+]$ is called the range of the optimal values of the ILP model (2). Computing of the best value z_{opt}^- is easy, but obtaining the worst value z_{opt}^+ is much more involved [3].

The model $\min\{c^T x | Ax = b, x \geq 0\}$, where $c \in [c^-, c^+] \subseteq \mathbb{R}^n, b \in [b^-, b^+] \subseteq \mathbb{R}^m, A \in [A^-, A^+] \subseteq \mathbb{R}^{m \times n}$ is called B-stable with basis B, if B has been an optimal basis for any characteristic model. B-stability (basis stability) conditions are as [5]: 1. A_B is non-singular or regular (regularity), 2. $A_B^{-1}\mathbf{b} \geq 0$, (feasibility), 3. For minimization problem if $\mathbf{c}_N^T - \mathbf{c}_B^T A_B^{-1} A_N \geq 0^T$, then A_B is optimal (optimality). Under the assumption of basis stability, it is possible to obtain the optimal solution set of ILP.

Let Y_m be the set of all $\{-1, 1\}$ m-dimensional vectors, i.e. $Y_m = \{y \in \mathbb{R}^m | |y| = e\}$, where e=$(1, 1, ..., 1)^T$. For a given $y \in \{-1, 1\}^T$, by $T^y = diag(y_1, y_2, ..., y_m)$, we denote the corresponding diagonal matrix. For each $x \in \mathbb{R}^n$, we denote its sign vector sign(x) by,

$$sign(x_i) = \begin{cases} 1 & if \quad x_i \geq 0 \\ -1 & if \quad x_i < 0, \end{cases}$$

where, $i = 1, 2, ..., n$.

Then we have $|z| = T_z x$, where $z = sign(x) \in Y_n$. Let an interval matrix $A^{\pm} = [A^-, A^+] = [A^c - \Delta_{A^{\pm}}, A^c + \Delta_{A^{\pm}}]$ and an interval vector $b^{\pm} = [b^-, b^+] = [b^c - \Delta_{b^{\pm}}, b^c + \Delta_{b^{\pm}}]$. For each $y \in Y^m$ and $z \in Y^n$, A_{yz} and b_y are defined as follows [5]:

$\mathrm{A}_{yz} = A^c - T_y \Delta_{A^{\pm}} T_z, b_y = b^c + b_y \Delta_{b^{\pm}}$.

[3] The best and the worst values of the objective function of the ILP model (1) could be obtained as follows:

$$z_{opt}^- = \{c^{-T} x | A^- x \leq b^+, A^+ x \geq b^-, x \geq 0\},$$

$$z_{opt}^+ = sup_{y \in Y_m} f(A_{ye}, b_y, c^+),$$

where, f(A, b, c)=inf$\{c^T x | Ax = b, x \geq 0\}$.

[3] Computing the worst value of the objective function of the ILP model (1) is NP-hard.

III. Overview of genetic algorithm

The genetic algorithm (GA) is a stochastic global search technique that solves problems by imitating processes observed during natural evolution [6]. GA adopts the spirit species extinction and also finds the approximate optimal solution after the process of coding, decoding and constant operation (selection, crossover and mutation). The procedure genetic algorithm is implemented as follows:

Choose initial population
Evaluate the fitness of each individual in the population
Repeat
    Select best-ranking individuals to reproduce
    Breed new generation through crossover and mutation (genetic operations) and give birth to offspring
    Evaluate the individual fitnesses of the offspring
    Replace worst ranked part of population with offspring
Until <terminating condition>.

IV. Analysis of validity

Suppose the intervals $z_{opt} = [z_{opt}^-, z_{opt}^+]$ and $z = [z^-, z^+]$ as the exact range and the obtained range by GA of the optimal values of the ILP model (1), respectively. So, we define the degree of uncertainty for the obtained range as follows:

$$UD(z) = \left(\frac{|z^- - z_{opt}^-| + |z^+ - z_{opt}^+|}{|z_{opt}^+ - z_{opt}^-|} \right) \times 100.$$

This indicator clearly shows that the obtained range is closer to the exact optimal range, if the degree of uncertainty associated with it, is close to be zero percent. As well as, the degree of feasibility for the obtained range is defined as follows:

$$FD(z) = \left(\frac{|(z_{opt}^+ - z_{opt}^-) - (z^+ - z^-)|}{|z_{opt}^+ - z^+| + |z_{opt}^- - z^-|} \right) \times 100.$$

Also, this indicator clearly shows that the obtained range is sufficiently feasible and logical, if the degree of feasibility corresponding with it, is close to one hundred percent. So, the obtained range z is completely valid, if $UD(z) = 0\%$ and $FD(z) = 100\%$.

V. Illustrative example and comparison of the results

In this section, we determine the optimal values for the following example with help GA method and then we compare the obtained optimal values through it with the obtained optimal values through Monte Carlo simulation based on three distribution functions, normal, uniform and beta [1] and the exact optimal values. Moreover, to show validity of GA method, we use the indicators of the degree of uncertainty and the degree of feasibility.

Consider the interval linear programming

$$\begin{aligned} &\text{Min } \ Z^{\pm} = x_1^{\pm} + x_2^{\pm} \\ &\text{s.t.} \\ &\qquad [2, 3] x_1^{\pm} + x_2^{\pm} = [3, 4] \\ &\qquad [4, 6] x_1^{\pm} + [-3, -2] x_2^{\pm} = [1, 5] \\ &\qquad x_1^{\pm}, x_2^{\pm} \geq 0. \end{aligned} \tag{3}$$

By theorem II, the best and the worst values of the objective function of ILP model (V) are $z^{-}_{opt} = 1$ and $z^{+}_{opt} = 3.1$, respectively. So, the range of exact optimal values of the objective function is $[1, 3.1]$. Also, the obtained optimal values through Monte Carlo simulation based on three distribution functions, normal, uniform and beta for Example (V), are persented in Table (I) [1].

TABLE I
RESULTS OF MONTE CARLO SIMULATION BASED ON THREE DISTRIBUTIONS

Distribution	***Values of z***	$UD(z)$	$FD(z)$
Normal	***[1.0394, 2.8585]***	***13.38 %***	***100%***
Uniform	***[0.9512, 2.9398]***	***9.95%***	***53.30%***
Beta	***[0.9384, 2.8186]***	***16.33%***	***64.08%***

The initial parameters related to GA method for solving this ILP problem are considered in Table (II). The obtained optimal values for Example (V), through GA method are presented in Table (III). Also, the values of z resulted from it, are shown in Figures 1 and 2.

TABLE II
INITIAL PARAMETERS RELATED TO GA METHOD

Method	***GA***
parameter	***Population: 20, Crossover rate: 10*** ***Mutation rate: 0.02, Generation: 100***

TABLE III
RESULTS OF GA METHOD

Method	***Values of z***	$UD(z)$	$FD(z)$
GA	***[1.004, 3.09]***	***0.67 %***	***100 %***

Fig. 1. The best value of z resulted from GA method

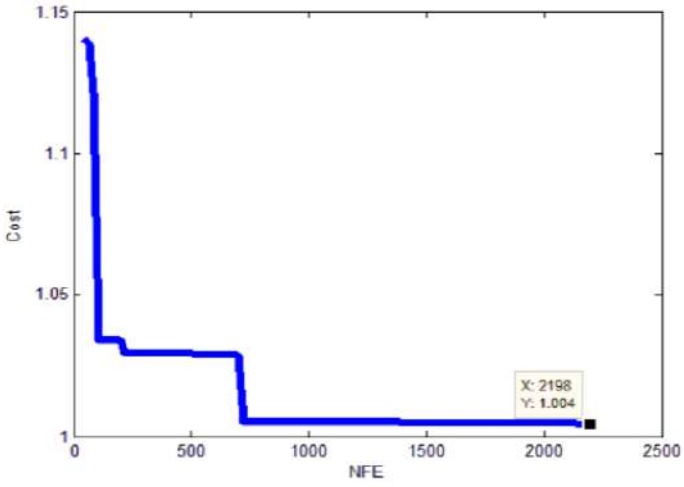
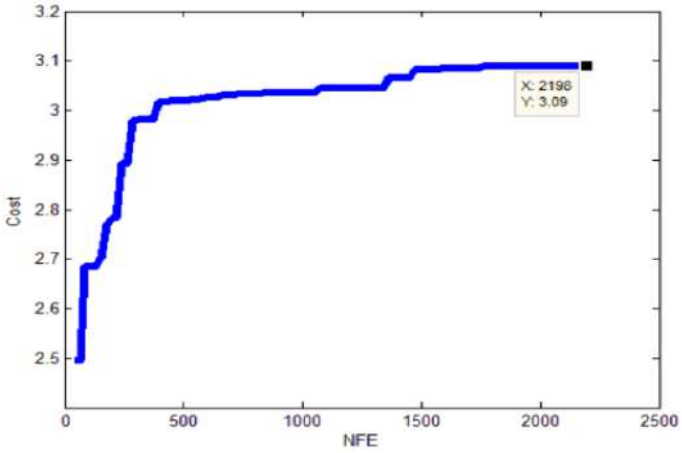


Fig. 2. The worst value of z resulted from GA method



VI. CONCLUSION

Computing the worst value of the objective function is very complicated for ILP models with equality constraints. GA has been used to explore the values of the objective function. The obtained results in Example (V), show that the obtained optimal values, through GA are much better than from Monte Carlo simulation results. Moreover, the obtained optimal values, through GA are much accurate and completely logical, while the obtained optimal values, through Monte Carlo simulation are highly inexact and as well as a large part of them are illogical and impossible. So, it is clear that obtaining optimal range by GA is not only possible but also by increasing the number of generation in its implementation, the exact optimal range can be obtained.

REFERENCES

[1] M. Allahdadi, A. Khaje Golestane, (2015), Monte Carlo Simulation for Computing the Worst Value of the Objective Function in the Interval Linear Programming. Int. J. Appl. Comput. Math. pp 1–10, Springer, India.

[2] M. Allahdadi, H. Mishmast Nehi, H. A. Ashayerinasab, M. Javanmard, (2016), Improving the modified interval linear programming method by new techniques, Information Sciences, 339, pp. 224–236.

[3] M. Fiedler, J. Nedoma, J. Ramik, K. Zimmermann, (2003), Linear Optimization Problems With Inexact Data, New York.

[4] M. Hladik, (2014), How to determine basis stability in interval linear programming, Optim. Lett., 8, 375–389.

[5] M. Hladik, (2012), Complexity of necessary efficiency in interval linear programming and multiobjective linear programming. Optim. Lett. 6, 893–899.

[6] N. Khilwani, A. Prakash, R. Shankar, M. K. Tiwari, (2008), Fast clonal algorithm, Eng. Appl. Artif. Intell., 21, 106–128.

[7] Lai, K.K., Wang, S.Y., Xu, J.P., Zhu, S.S., Fang, Y. (2002), A class of linear interval programming problems and its application to portfolio selection, IEEE Trans. Fuzzy Syst., 10(6), 698–704.

[8] C.Y. Lin, P. Hajela, (1992), Genetic Algorithms in Optimization Problems with Discrete and Integer Design Variables, Eng. Optimization, Vol. 19, 309-327.

[9] S. C. Tong, (1994), Interval number, fuzzy number linear programming, Fuzzy Sets Syst., 66, 301–306.

[10] S. J. Wu, P. T. Chow, (1995), Genetic Algorithms for Nonlinear Mixed Discrete-Integer Optimization Problems via Meta-Genetic Parameter Optimization, Eng. Optimization, Vol. 24, 137–159.

[11] M. Javanmard, H. Mishmast Nehi, Rankings and operations for interval type-2 fuzzy numbers: a review and some new methods, Journal of Applied Mathematics and Computing, DOI: https://doi.org/10.1007/s12190-018-1193-9, First Online: 31 May 2018.

[12] M. Javanmard, H. Mishmast Nehi, Arithmetic operations and ranking methods for interval type-2 fuzzy numbers; A review and some new methods, 9th International Iranian Operation Research conference, IORC, Shiraz, Iran (2016), 928-933.

[13] M. Javanmard, H. Mishmast Nehi, Solving the interval type-2 fuzzy linear programming problem by the nearest interval approximation, 6th Iranian Joint Congress on Fuzzy and Intelligent Systems, CFIS, Kerman, Iran, DOI: https://doi.org/10.1109/CFIS.2018.8336609, Publisher IEEE, Feb 2018.

[14] M. Javanmard, H. Mishmast Nehi, A Solving Method for Fuzzy Linear Programming Problem with Interval Type-2 Fuzzy Numbers, Int. J. Fuzzy Syst., DOI: https://doi.org/10.1007/s40815-018-0591-3, Published Online: 17 January 2019.

[15] M. Javanmard, H. Mishmast Nehi, Interval type-2 triangular fuzzy numbers; New ranking method and evaluation of some reasonable properties on it, 5th Iranian Joint Congress on Fuzzy and Intelligent Systems, CFIS, **8003587**(2017), 4–6.

[16] M. Javanmard, H. Mishmast Nehi, Solving interval type-2 fuzzy linear programming problem with a new ranking function method, 5th Iranian Joint Congress on Fuzzy and Intelligent Systems, CFIS, **8003586**(2017), 1–3.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)
March 2-4, 2022, Higher Education Complex of Bam

Detection and visualization of COVID-19 in chest X-ray images using CNN and Grad-CAM (GCCN)

Hasan Hammad, Hassan Khotanlou
Department of Computer Engineering, Bu Ali Sina University, Hamedan, Iran,
Hasanhammad13@gamil.com, khotanlou@basu.ac.ir

***Abstract*— The quick and early detection of COVID-19 infection is very important in the fight against the pandemic. Deep learning can be considered as a helpful method to provide help and assist the medical staff to detect the infection of COVID-19, which will definitely have a positive effect in controlling the outbreak of COVID-19. In this paper, we will propose a simple CNN based deep learning model called Grad-CAM CNN (GCNN), for the purpose of detecting the infection of COVID-19 disease from chest X-ray images and visualizing a heat map with the help of the Grad-CAM technique in order to determine which area in the X-ray image of the chest has COVID-19. Since CNN is a very powerful method in processing images, we use it to build the model. We evaluated the performance of the proposed method on public online available datasets of X-ray images of the chest from Kaggle. The proposed method is able to achieve an accuracy score of 97.78 % using a learning rate of 0.003 with the Adam optimizer. In the light of the promising results obtained from this method, it is possible to say that the proposed method can be helpful in the early diagnosis in the upcoming waves of COVID-19.**

***Index Terms* —GCNN; Deep Learning; Grad-CAM; Convolutional Neural Network (CNN); Medical Images; COVID 19; Chest X-ray**

I. INTRODUCTION

Since COVID-19 is a serious disease, earlier detection is important in the upcoming waves. The current testing method has some limitations in processing time and accuracy. So, aiming to provide help and quick healthcare to the infected patients, there is a pressing need for early and quick detection of infection. Studies have proven that we can consider X-ray images and computed tomography scans of the chest as diagnosis methods [1]. The advantage of imaging is that it is a cheaper and simple technique, as well it has good sensitivity, and a fast turnaround time. But, using chest X-ray images instead of computed tomography scans, to detect the infection, has a major advantage as it is cheaper and more available, especially for the countries where resources are scarce [2] and test kits are not sufficient. Moreover, some studies have proven that radiographic abnormalities of COVID-19 positive cases could be presented in chest X-ray images [3].

The generated diagnosis reports and medical images lead to a rapid increase in the medical data. The way that we use these data has changed by deep learning. We call it deep learning because the network consists of many layers with a large number of trainable parameters. Recurrent Neural Network (RNN), Convolutional Neural Network (CNN) and Artificial Neural Network (ANN) are the three major methodologies in deep learning. The two types which are RNN and ANN can be used with numbers or text data whereas, CNN can be used with images [4].

CNN and the human brain are very similar in the way of process images. The entire visual field in the human brain is covered with a network consisting of a large number of neurons. CNN processes the images in a very similar way [5]. In this paper, we used a simple CNN-based deep learning model and the Grad-CAM technique. Aiming to detect the infection of COVID-19 disease, we constructed this model to classify the input image after processing and extracting features from the input X-ray image of the chest. In the field of medical images processing, high accuracy machine learning models are needed in order to achieve an early and fast diagnosis. Therefore, we used the Adam optimizer in the model and changed many parameters, to be able to make the performance and the classification and validation accuracy of the proposed model better. This work revolves around implementing a simple convolutional neural network model consisting of 11 layers,

that is basically trained to be able to detect the infection of COVID-19 disease, from the given X-ray images of the chest. We also used the Grad-CAM technique for the purpose of determining which zone in the chest has COVID-19 which in turn will make the proposed method more explainable and easier to understand.

We collected the data from publicly online available datasets of COVID-19 and normal X-ray images of the chest from Kaggle. We used the combined dataset, which contains two classes COVID-19 and Normal, to train, validate and test the model.

The aim of this work is to propose a simple method that can give quick and good results and can be used on a huge number of patients to assist doctors and radiologists to detect the infection of COVID-19 in a faster and more precise way.

II. Related Works

Recently, the diagnosis of COVID-19 infection has pulled the attention of many researchers all around the world. In the last few years, in order to develop new deep learning models, algorithms and computer-aided systems, a lot of studies and works were done. Many of these studies revolved around using the popular deep learning methodologies.

The COVIDX-Net proposed by Hemdan et al. [6], is a deep learning model that can detect the infection of COVID-19 from X-ray images of the chest. The performance of InceptionResNetV2, InceptionV3, VGG19, MobileNetV2, ResNetV2, DenseNet201 and Xception, was analyzed in their study. Their dataset included 25 COVID-19 positive X-ray images of the chest and 25 normal X-ray images of the chest. Their results showed that the two CNN models VGG19 and DenseNet were able to achieve the highest scores with a 90.00% of accuracy. Wang et al [3] proposed the COVIDNet deep learning model. Aiming to detect the infection of COVID-19 from the given X-ray images of the chest, the 1 * 1 convolutional deep learning technique was applied to the dataset. They used a combined dataset, for a multi-classification of COVID-19, normal and non-COVID-19. Their dataset is composed of Dr. Joseph Cohen's COVID-19 X-ray images of the chest [7], dataset of X-ray images of the chest from Kaggle [8] and dataset of COVID-19 X-ray images of the chest from GitHub [9]. Their test accuracy score was 92.6%. In another study, Maghdid et al [10], used a transfer learning approach and deep learning-based model in order to automatically detect the infection of COVID-19 using convolutional neural networks. They were able to achieve a 94.00% of accuracy. Singh et al [11], proposed a deep learning model based on CNN. They inspired their proposed model from the architecture of DarkCovidNet's [12]. Our proposed model is composed of 11 layers whereas their proposed model is composed of 19 layers. They achieved a classification accuracy of 87.00%. In the study of Ghoshal et al [13], aiming to improve the human-machine diagnostic performance, they investigated how the uncertainty in deep learning solutions can be estimated by Dropweights based Bayesian Convolutional Neural Networks. To be able to overcome the small number of images in their dataset, real-time augmentation techniques and a transfer learning approach were used. Their proposed method, which used VGG16 deep learning model, was able to provide a 92.86% of accuracy score.

As we mentioned before, many works have been done aiming to detect the infection of COVID-19, many of which provided very good accuracy scores. Here we are trying to keep on by proposing a new CNN-based deep learning model with acceptable accuracy for the purpose of detecting the infection of COVID-19 from X-ray images of the chest.

III. Proposed Method

Our proposed method (GCNN) consists of a simple CNN-based deep learning model equipped with the Grad-CAM technique. Intending to detect the infection of COVID-19 from the given chest X-ray images, the simple CNN-based model was used. The Grad-CAM technique returns a heat map, which is what the model has learned in the last convolutional layer, and which in turn will be mapped on the given image, to determine which area in the chest has COVID-19 infection.

A. Pre-processing

After collecting the dataset, we pre-processed and implemented the data augmentation operations using ImageDataGenerator from Keras library [14]. The augmentation operations are listed as follows:
Scaling: we rescaled the images with a rescaling factor of 1/255 to normalize the images. Zooming: zooming rate of 0.2 was used, which will implement randomly zooming inside the input images between 80% (zoom in) and 120% (zoom out). Horizontal flipping: we set the horizontal flipping to true in order to randomly flip the input images horizontally. After that, we partitioned the dataset into three sets training (80%), testing (20%), validation (20% of the training set).

B. Deep Learning Model

The proposed model is composed of several layers (see Fig. 1 and Fig. 2). The purpose of using convolutional layers is to extract features from the given images. We used three 2D convolutional layers. The first 2D convolutional layer, which is the first layer in the proposed model, has an input shape of (150 * 150 * 3). Here the kernel size is 5 * 5 and the number of filters is 32. In the second and third 2D convolutional layers, we used the same kernel size with 64 filters. The rectified linear unit (ReLU) was used in all the convolutional layers. We used two pooling layers with a pool size of 2 * 2 in order to reduce the features in the CNN architecture. The max-pooling layers, divide the input image into rectangular pools, then calculate the maximum of each rectangle pool. Moreover, these max-pooling layers reduce the overfitting, as a result of this, the model generalization will be improved. The three Dropout layers were used to randomly ignore some hidden and visible units, which help to prevent overfitting. The three dropout layers drop 50% of the units. Two fully connected layers were used in this model. We used the rectified linear unit (ReLU) in the first one, which is the 8th layer with 256 units and in the second layer,

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 150, 150, 32)	2432
max_pooling2d (MaxPooling2D)	(None, 75, 75, 32)	0
dropout (Dropout)	(None, 75, 75, 32)	0
conv2d_1 (Conv2D)	(None, 75, 75, 64)	51264
conv2d_2 (Conv2D)	(None, 75, 75, 64)	102464
max_pooling2d_1 (MaxPooling2D)	(None, 37, 37, 64)	0
dropout_1 (Dropout)	(None, 37, 37, 64)	0
flatten (Flatten)	(None, 87616)	0
dense (Dense)	(None, 256)	22429952
dropout_2 (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 1)	257

Total params: 22,586,369
Trainable params: 22,586,369
Non-trainable params: 0

Figure 1. Summary of the proposed CNN model

the Sigmoid activation function was used with 1 unit, since it is the last layer in this model.

We trained the proposed model on the dataset using the Adam optimizer for 8 epochs. We took the learning rate as 0.003 with a batch size of 32. In Table I, more details about the proposed CNN model are provided.

C. Grad-CAM Technique

The Gradient Weighted Class Activation Mapping (Grad-CAM) technique was provided by Selvaraju et al. [15]. The Grad-CAM technique provides a more understandable view of deep learning models. Furthermore, it can be very helpful in detection works because it shows more about the model. In this paper, we made use of the Grad-CAM technique to generate a color visualization of the CNN model output, so that it can be more understandable. As shown in Fig. 3, the CNN model was used to detect the infection of COVID-19. After determining the predicated labels, the Grad-CAM technique was applied to the last convolutional layer.

IV. EVALUATION

A. Dataset

Collecting a large dataset of medical images is quite challenging because it requires multiple expert opinions to overcome the human error. Furthermore, it requires a lot of time from the medical experts to collect the data. To be able to overcome the previous issues, we used two different sources to collect the dataset. The first is a publicly online available dataset from Kaggle [16]. This dataset is split into two directories train and test, both contain three subdirectories NORMAL, COVID19 and PNEUMONIA. We removed the PNEUMONIA subdirectory from both train and test directories since we are building a model for binary classification. Now the dataset contains a total of 2159 X-ray images of the chest. In Fig. 4 sample images from this dataset are provided. The second dataset is the COVIDx dataset structured by Wang et al. [17], from which we took 1007 COVID-19 infected X-ray images of the chest, in order to balance the data in the first dataset. In Fig. 5 sample images from this dataset are provided. The distribution of images according to their source dataset is shown in Table II. The total dataset has 3166 images. We partitioned the dataset into three sets training, testing, and validation. More precisely, 20% of the dataset was used to test the model and 80% for training but, in the training stage, the training set was partitioned into two sets: 80% was used to train the model whereas the remaining 20% was used for validation. The numbers of images in each set are provided in Table III.

B. Performance Measures

For the purpose of measuring the performance of the proposed method, we used the following matrices:

1) Accuracy: is the ratio between the summation of true positive and true negative images, to the total number of images.

$$Accuracy = \frac{TP + TN}{FP + FN + TP + TN}$$

2) Sensitivity: is the matrix that evaluates the model's ability to predict true positive images of each available category.

$$Sensitivity = \frac{TP}{TP + FN}$$

3) Precision: is the ratio between the number of true positive images correctly classified to the summation of images classified as positive.

$$Precision = \frac{TP}{TP + FP}$$

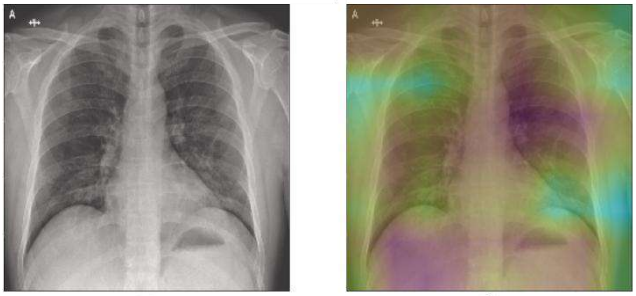
4) Specificity: is the matrix that evaluates the model's ability to predict images classified as true negative of each available category.

$$Specificity = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

Figure 2. Architecture of the proposed CNN model

TABLE I. PROPOSED METHOD PARAMETERS DETAILS.

Method	Proposed Method Parameters									
	Optimizer function	***Output layer classifier***	***Size of max pooling kernels***	***Dropout layer size***	***Number of epochs***	***Batch size***	***Number of fully connected layers neurons***	**Activator function**	***Size of convolution kernels***	***Number of convolutional layers kernels***
GCNN	Adam	Sigmoid	2*2	0.5	8	32	256 and 1	ReLU	5*5	32 and 64



(a) (b)

Figure 3. (a) Original COVID-19 positive chest X-ray image (b) The heat map of the original COVID-19 positive chest x-ray image.

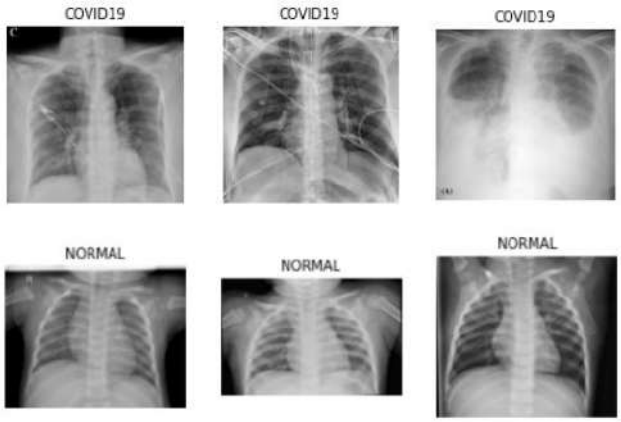


Figure 4. Sample images from Dataset-1 [16]

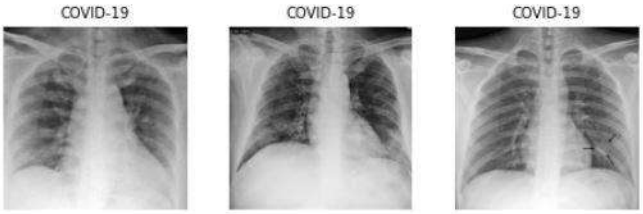


Figure 5. Sample of COVID-19-positive images from Dataset-2 [17]

TABLE II. X-RAY IMAGES DISTRIBUTION ACCORDING TO THEIR SOURCES.

Dataset	Images Classes in The Dataset	
	COVID-19	***Normal***
Dataset-1[16]	576	1583
Dataset-2[17]	1007	-

TABLE III. SAMPLES DISTRIBUTION IN THE DATASET

Set	Images Classes in The Dataset	
	COVID-19	***Normal***
Training	1014	1013
Validation	253	253
Test	316	317
Total	1583	1583

C. Results and Discussion

Our proposed method (GCNN) and some state-of-the-art methods are compared according to accuracy, sensitivity and specificity. Some studies did not report their sensitivity and specificity. Furthermore, some related studies which achieved good scores, have a lack in the number of the COVID-19 positive X-ray images of the chest. Table IV shows the performance comparison with some binary classification studies based on X-ray images. Sethy et al. [18], proposed a hybrid method consisting of support vector machine algorithm and deep learning methods. For binary classification, their achieved accuracy score was 95.38%. Hemdan et al. [6] and Narin et al. [19], used transfer learning, end-to-end learning algorithms and deep features extraction for the purpose of detecting the infection of COVID-19 using X-ray images of the chest. Panwar et al. [20], proposed the nCOVnet, which is based on deep learning. For binary classification, their achieved

TABLE IV. THE PERFORMANCE COMPARISON WITH SOME BINARY CLASSIFICATION STUDIES.

Study	Method's Details and Performance Scores				
	Number of COVID-19 Images	***Method***	***Acc (%)***	***Se (%)***	***Sp (%)***
Hemdan et al. [6]	25	VGG19, ResNet, DenseNet, Inception, Xception	90%	100%	80%
Sethy et al. [18]	127	Deep features, ResNet-50, SVM	95.38%	-	-
Narin et al. [19]	341	InceptionV3, ResNet50, ResNet101, ResNet152, Inception-ResNetV2	96.1%	-	-
Panwar et al. [20]	142	nCOVnet, deep CNN, transfer learning	88.10%	97.62%	78.57%
Sharifrazi et al. [21]	77	2D-CNN (CNN-SVM + Sobel)	99.02%	100%	95.23%
Proposed Method	1583	GCNN, Deep CNN	97.78 %	98.42 %	97.15 %

accuracy was 88.10%. Sharifrazi et al. [21], implemented a fusion of support vector machine, CNN and Sobel filter, in order to detect the infection of COVID-19 from X-ray images of the chest. The highest accuracy, specificity and sensitivity achieved by their method were 99.02%, 95.23%, 100%. Hemdan et al. [9], mentioned that their main problem in the work was the small number of COVID-19 positive X-ray images. Narin et al. [19], mentioned that the high accuracy was obtained because of the large number of normal images. Their achieved accuracy cannot be very accurate considering that there are few differences between the COVID-19 infected X-ray images and the normal X-ray images of the chest. Whereas the composite dataset that we used in this work was balanced and included a large number of Normal and COVID-19 positive X-ray images of the chest. The accuracy, sensitivity, precision and specificity achieved by this method are 97.78%, 98.42%, 97.19% and 97.15%, considering that the proposed method is simpler than the mentioned studies.

The proposed method is also compared with two pre-trained methods, VGG16 and ImageNets. We performed the evaluation on the same dataset. The achieved scores are given in Table V. The topmost result was obtained by MobileNets with an accuracy score of 98.73%. The second-best result was obtained by the proposed method (GCNN) with an overall accuracy of 97.78%. Even though, The MobileNets and VGG16 models were implemented using transfer learning. Moreover, MobileNets and VGG16 models are larger and deeper than the proposed method. The proposed method was able to achieve better results than the VGG16 pre-trained model. For now, we cannot claim that our proposed method (GCNN) is completely better than the other methods because of the fact that a large, clear and organized dataset is not available yet. In the future, we aim to test the performance of the (GCNN) on a challenging dataset.

V. CONCLUSION

The proposed method, which is based on a simple CNN deep learning model was able to detect the infection of COVID-19 with a 97.78% of accuracy in a binary classification of normal and COVID-19 positive X-ray images of the chest. The main limitation in our work was the insufficient number of high quality COVID-19 positive X-ray images of the chest. we used a composite dataset to overcome this limitation. On the brighter side, our proposed model achieved a very good accuracy, which is going to have a significant role in improving the detection of COVID-19 infection. Later on, aiming to improve the performance of the proposed method, we want to extend our work and train our model on a larger and more challenging dataset, as well as develop a multiclass classification model. In the meantime, this study and many others provide many methods with high accuracy which will provide great support to the medical staff in the going COVID-19 pandemic. We hope that our work inspires others to improve their works and help the medical staff and the community.

TABLE V. THE PERFORMANCE SCORES OF THE PROPOSED METHOD (GCNN) AND THE PRE-TRAINED METHODS

Method	**Performance Scores**		
	Accuracy (%)	***Sensitivity (%)***	***Specificity (%)***
VGG16	96.99%	93.98%	93.98%
MobileNets	98.73%	99.68%	97.78%
Proposed Method (GCNN)	97.78 %	98.42%	97.15%

REFERENCES

[1] D. Wang et al. "Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China," JAMA, vol. 17, no. 11, March 2020, doi: 10.1001/jama.2020.1585.

[2] G. D. Rubin et al. "The Role of Chest Imaging in Patient Management During the COVID-19 Pandemic: A Multinational Consensus Statement

From the Fleischner Society," chest, vol. 158, no. 1, April 2020, doi: 0.1016/j.chest.2020.04.003.

[3] L. Wang, Z. Q. Lin, and A. Wong, "COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images," Sci Rep, vol. 10, no. 1, November 2020, doi: 10.1038/s41598-020-76550-z.

[4] M. I. Razzak, S. Naz, and A. Zaib, "Deep Learning for Medical Image Processing: Overview, Challenges and Future," arxiv.org/abs/1704.06825v1.

[5] Towardsdatascience.com, "Convolutional Neural Networks, Explained," August 2020, https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939, [Accessed: 14 - November- 2021].

[6] E. E. D. Hemdan, M. A. Shouman, and M. E. Karar, "COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images," arxiv.org/abs/2003.11055.

[7] J. P. Cohen, P. Morrison, and L. Dao, "COVID-19 Image Data Collection,"arxiv.org/abs/2003.11597v1.

[8] Kaggle, Kaggle's Chest X-Ray Images (Pneumonia) dataset, https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/.

[9] Figure1-COVID-chestxray-dataset, https://github.com/agchung/Figure1-COVID-chestxray-dataset.

[10] H. S. Maghdid, A. T. Asaad, K. Z. Ghafoor, A. S. Sadiq and M. K. Khan, "Diagnosing COVID-19 Pneumonia from X-Ray and CT Images using Deep Learning and Transfer Learning Algorithms," arxiv.org/abs/2004.00038.

[11] S. Singh, P. Sapra, A. Garg and D. K. Vishwakarma, "CNN based Covid-aid: Covid 19 Detection using Chest X-ray," 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), 2021, pp. 1791-1797, doi: 10.1109/ICCMC51019.2021.9418407.

[12] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim and U. R. Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," Comput Biol Med. 2020 Jun;121:103792. doi: 10.1016/j.compbiomed.2020.103792. Epub 2020 Apr 28. PMID: 32568675; PMCID: PMC7187882.

[13] B. Ghoshal and A. Tucker, "Estimating Uncertainty and Interpretability in Deep Learning for Coronavirus (COVID-19) Detection," arxiv.org/abs/2003.10769.

[14] Tensorflow, "Keras, preprocessing, image, ImageDataGenerator," https://www.tensorflow.org/api_docs/python/tf/keras/preprocessing/image/ImageDataGenerator.

[15] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 618-626, doi: 10.1109/ICCV.2017.74.

[16] Kaggle, Kaggle's Chest X-ray (Covid-19 & Pneumonia), https://www.kaggle.com/prashant268/chest-xray-covid19-pneumonia.

[17] COVIDx Dataset, https://github.com/lindawangg/COVID-Net/blob/master/docs/COVIDx.md, (2021).

[18] P. K. Sethy, S. K. Behera, P. K. Ratha, P. Biswas, "Detection of Coronavirus Disease (COVID-19) Based on Deep Features and Support Vector Machine," Preprints 2020, 2020030300.

[19] A. Narin, C. Kaya and Z. Pamuk, "Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks," (2020).

[20] H. Panwar, P. K. Gupta, M. K. Siddiqui, R. Morales-Menendez and V. Singh, "Application of deep learning for fast detection of COVID-19 in X-Rays using nCOVnet," Vol. 138, 2020, ISSN 0960-0779, https://doi.org/10.1016/j.chaos.2020.109944.00.

[21] D. Sharifrazi et al. "Fusion of convolution neural network, support vector machine and Sobel filter for accurate detection of COVID-19 patients using X-ray images," arxiv.org/abs/2102.06883v1.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Detection of multiple emotions in texts using a new deep convolutional neural network

Habib Izadkhah

Department of Computer Science, Faculty of Mathematics, Statistics and Computer Science, University of Tabriz, Tabriz, Iran, izadkhah@tabrizu.ac.ir

***Abstract*— Identifying emotions from the texts can be used in almost every aspect of our daily lives, such as improving computer-human interactions, monitoring people's mental health, or modifying/improving business strategies based on customers' emotions. Deep learning techniques have performed well compared to other machine learning methods in all learning problems. All existing machine learning methods for recognizing emotions are taught on datasets that includes single-emotional and multi-emotional samples. Our observations of working with these techniques show that these systems tend to learn more single-emotion samples than multi-emotion samples in a dataset. We also looked at a large number of texts and found that the number of texts from which only one emotion can be deduced is very small compared to texts from which more than one emotion can be deduced. Therefore, in general, the accuracy of existing methods is low. To deal with these two limitations, in this paper, we first created a dataset using all available data so that all texts have at least two emotions. To improve accuracy, because the convolutional neural network (CNN) has performed so well in image processing, we have then proposed a new CNN architecture for extracting emotions from texts. To find semantic, syntactic, and word similarity, we have used *fast text* and *GloVe* to embed text data into numeric representation. We have also improved the accuracy of the proposed architecture by using the *attention* property. The results demonstrated that the proposed model is more accurate compared to the existing methods.**

***Index Terms*— Natural language processing (NLP), Deep learning, Convolutional neural network, Emotions**

I. Introduction

Emotions reflect people's behavior and thoughts. Expressing emotions about different events, people, environment and every little other thing is always done in human daily life. The common ways to express feelings to others is through writing, speech and facial expressions. Nowadays, text is the most common form of communication on social networking platforms. People express their feelings and emotions through posts on social networks. Now, people can exchange facts, opinions, emotions, feelings, and even the intensity of emotions on a variety of topics in short texts. The analysis of emotions and feelings expressed in the content of social networks has attracted researchers in the field of natural language processing research [1].

Accurate knowledge of human emotions is a difficult task due to the versatility and ambiguities in emotions and feelings. Recognizing and analyzing emotions from the text is relatively more complex than recognizing feelings. Although the terms emotions and feelings are considered synonymous, they represent two different concepts in text analysis. The American Psychological Association defines emotions as a "complex pattern of change, including physiological arousal, cognitive processes, and behavioral responses, in response to a situation that is considered personally significant". Feelings are generally defined as being influenced by emotions. Happiness, anger and love are examples of emotions, and positive, negative and neutral are examples of feelings. For example, a person whose posts show sadness, frustration, and depression can be identified immediately and referred to a specialist before an unpleasant event such as suicide [2]. As another example, consider an employee sending a violent email to a co-worker or boss. A tool that can analyze the emotions of the email before sending an email response informs the employee of the person's immorality, which is very useful to protect the employee. Other useful tools that can be used to recognize emotions from the text include recommender systems that aim to personalize recommendations based on users' emotions [3].

Emotion classification can be divided into two main groups: single-labeled emotions and multi-labeled emotions. In multi-label emotions, each text can be associated with a set of emotions.

The purpose of this paper is to analyze multi-labeled emotions in texts. Traditional machine learning methods are not suitable for this purpose because they require feature engineering and this process is a time consuming process and requires background knowledge. Since the use of deep learning in recent years has shown acceptable results in most areas, therefore, this paper is the development of a system using deep learning that solves the problem of multi-labeled emotion analysis. To this end, we first create a dataset with all the multi-labeled samples of the combination of the existing datasets, and we then present a new architecture using convolutional

layers.

The contributions of this paper can be summarized as follows:

1) The first work to identify at least two emotions from the text. Our research on emotion identification datasets, SemEval-2018, BMET, CBET, NLPCC2018, and REN-CECPS, shows that in most texts there are at least two emotions and there are very few texts that show only one emotion. On the other hand, existing machine learning-based methods for identifying emotions learn single emotions better than multi-emotions, and therefore respond more to single-emotion recognition, and this reduces the accuracy of these methods.
2) Proposing a new deep convolutional neural network. In this paper, we designed a module, called X-module, that considers both the depth and width of a network. This proposed architecture is unique in some ways, (1) we added the identity connection between the layers. This connection allows information is able to flow unimpeded throughout the entire network; (2) to overcome existing architecture limitations (i.e., computing overhead, memory consumption, and overfitting), we employed the idea of a wider network instead of just a deeper one, increasing the depth and width of the network; (3) an interesting feature of the proposed architecture is the use of two intermediate classifiers. This feature does not exist in other deep architectures for one-dimensional data. The intuition behind this addition is to prevent the "dying out" of the middle part of the network.

II. Related work

Existing works in the field of emotion analysis can be divided into two main approaches [4]. The first approach is the dictionary-based approach (the dictionary-based approach is essentially a rule-based approach). This approach calculates emotional score using a set of manually created rules as well as using a glossary. There are many dictionaries in this field, including WordNet-Affect and EmoSenticNet. These emotional glossaries include emotional keywords such as happiness, sadness, hatred, fear, and so on. Once a keyword is identified in the sentence, a label is assigned to the sentence. For example, if a sentence contains the keyword fear, it is related to the emotion category of fear. The second approach is the machine learning-based approach, which applies some machine learning algorithms to the training data set to be able to predict emotions in unobserved data. The second approach can also be divided into machine learning and deep learning approaches. In the following, we will have a review of past works in the field of emotion analysis.

In 2019, Mohamed Gabriel and Antonio Moreno [5] used a new deep learning approach to classify multi-labeled emotions on Twitter. They first proposed a way to turn the problem into a binary classification problem. Using this approach, a multi-label problem becomes one or more single-label problems, and in particular, the learned single-label classifications are applied.

In 2020, Seohui Park et al. [6] proposed a way to extract the emotions of a sentence using an emotion embedding model. To do this, they first created a word embedding model to create emotions using 144,701 collected tweet data containing the hashtag. They then extracted the representative emotional word in each sentence from the ROC story data. Then, they used the representative emotional word to classify the emotions of the sentence, which also uses cosine similarity.

In 2020, Sohasini and Badogo [7] used machine learning techniques to detect the emotions of Twitter messages. They used Bayes Naïve machine learning algorithms and the k nearest-neighbor algorithm to identify emotions in Twitter messages and then categorize Twitter messages into four emotional categories. They also conducted a comparative study of two supervied machine learning algorithms. They showed that the Bayes Naïve classifier performed better than the k nearest neighbor classifier.

In 2020, Jiavin Deng and Fuji Ren [8] identified multi-labeled emotions by discovering and extracting specific features and learning the correlation of features (emotions). They used a multi-label architecture, MEDA, to identify emotions and all related emotions in a text. MEDA consists mainly of two modules, a multi-functional feature extractor with emotion learners and emotion learners. This network consists of several channels by which they encode the characteristics of each emotion separately. The two-way GRU network is considered as a predictor of emotion sequence and emotions are predicted in a fixed direction, respectively. They also used a multi-label focal loss function to learn multi-label learning based on emotion correlation information.

III. Proposed Approach

This section describes the different parts of the proposed method for detecting multiple emotions from texts.

A. Dataset

A prerequisite for any machine learning system for automated learning is to have a set of data appropriate to that area. Since the purpose of this study is to work on datasets whose samples are completely multi-labeled (i.e., each sample of data is associated with at least two emotions), therefore, a dataset whose samples are all multi-labeled is needed. Because such a dataset does not exist, we proceed by combining all instances of the two reference datasets CBET [4] and semEval 18 [9] that relate to at least two emotions (or more than or equal to two classes). By combining multi-labeled samples of the two datasets, we were able to create a multi-labeled dataset for this study.

B. Data processing

A set of steps that, when performed on raw data, the dataset becomes an efficient format is called data processing. In general, datasets have flaws and defects that increase the error of the models. Therefore, data processing is an important requirement. In this paper, the following preprocessings are applied to the data.

Data Cleanup- Delete people usernames, site links, hashtags, spaces, and punctuation. These signals are not only useful, but

also cause noise and increase the computational load of the training network.

Deleting Stop Words- These words represent the most common words in any language. Examples of stop words are suffixes, pronouns, and conjunctions. 179 common stop words were identified and deleted. Deleting stop words helps models focus on more important words to make more accurate predictions. In addition, deleting these words reduces the input size and thus reduces the computational overhead.

Tokening- Simply put, tokening is the process by which text becomes a set of meaningful pieces called tokens. In this paper, we first identify the number of unique words and create a unique number for each word by creating a dictionary of words. Using this method, each sentence will be a vector of numbers. By doing this, the vectors are ready to feed into the word embedding layer.

C. Attention

Not all words in a sentence have the same contribution in representing the meaning of the sentence. To handle this case and extract words that have a great impact on the meaning of a sentence, we use the attention mechanism. This mechanism can be very helpful in analyzing emotions by paying attention to the meaning of words. To utilize the attention layer, as [10], we create a distributed representation of the entire document with respect to the highlighted words. To calculate the similarity of each word with the whole text, we multiply the dots of each word in the sequence with the final output of the encoding step. Then, the softmax function is used by passing the calculated similarity to it and draws the attention distribution to calculate the final layer yield.

D. Word embedding

In this layer, we used different approaches for the word embedding. First, we put the word embedding layer, like the other layers, in such a way that it learns the semantic vector of the words itself by training and giving weights to the words. In another approach, we used two pre-trained word embedding models, *fast text* and *GloVe*, to perform the role of both approaches, comparing and evaluating all three on the dataset being tested.

E. A new deep learning network architecture to identify emotions

Deep learning techniques have achieved remarkable results on various problems in recent years. One of the most popular and widely used deep learning techniques are convolutional neural networks (CNN), which have made great achievements in classifying 2D and 3D data such as images and videos. As mentioned, CNNs have gained much of their popularity in image classification and, in most cases, have been used with two-dimensional structures to classify images, for example, the diagnosis of medical images.

In this paper, a one-dimensional structure of convolutional neural networks is utilized to learn a one-dimensional dataset that detects the emotions. In the following, we design a new deep learning architecture using one-dimensional convolutional neural networks. The following three main reasons encouraged us to present a new architecture.

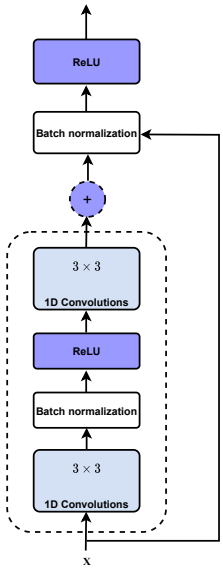


Fig. 1. A R-block.

1) For better learning as well as the efficiency of a neural network, it is necessary to have several layers and a large number of neurons in each layer. But the problem is that these deep networks are prone to overfitting, and it is also difficult to transfer gradient updates from the entire network.
2) Convolutional neural networks have the ability to perform feature extraction operations at different scales. Therefore, they have a better ability to learn. Naively stacking large convolution operations is computationally expensive. As mentioned earlier, these types of networks have been invented for visual data. We use them here for one-dimensional data.
3) Consideration of the Hebbian Principle — neurons that fire together, wire together.

Most state-of-the-art architectures for two-dimensional convolutional neural networks, e.g., ResNet50, DenseNet, and SE-Net, focus on the depth of the network, which increases the number of network parameters. Increasing the number of parameters will cause an overfiting. In this paper, for one-dimensional vectors, we design a module, called X-module, that considers both the depth and width of a network. To this end, inspired by residual block, we first design a block, called R-block, as illustrated in Fig. 1.

A principal feature in this block is to use of shortcut connections. In this block, using a shortcut connection, the input to the block is added to the output of the block. The block includes a stack of two 1D convolutional layers with the same number of filters so that the first convolutional layer is followed by a batch normalization layer and a ReLU activation function. Formally the block is formulated as the $y = f(x)+x$, where x is the input to the block, and $f(x)$ is a number of layers of 1D convolution, usually two of them or three in the

case of bottleneck layers. This results in the ability to train much deeper networks than what was previously possible.

The designed block is almost similar to the traditional convolutional neural networks. The most important modification is the *identity connection* between the layers. This connection allows information is able to flow unimpeded throughout the entire network. In Fig. 1, you can see the identity connection as the curved arrow originating from the input and sinking to the end of the residual block. This identity connection does not have any parameters and is just there to add the output from the previous layer to the layer ahead. Of course, it is possible to use, dropout, a 1×1 convolution or other convolutions instead of using a connection without a parameter.

In the following, we express the intuition behind this identity connection. Neural networks are used to estimate functions, and the estimation accuracy increases with increasing number of layers. Therefore, neural networks must be able to learn any simple or complex function. But published experimental results show that if we have deep enough networks, due to some problems such as the vanishing gradients and curse of dimensionality, the network may have difficulty learning simple functions such as the identity function. This is completely undesirable now. Therefore, the number of layers added is limited, which leads to a decrease in accuracy. It will also be observed that as the number of layers increases, the accuracy at one point becomes saturated and eventually disappears. This is usually not caused by overfitting. Therefore, it may seem that shallow networks learn better than their deeper counterparts, and this is completely counter-intuitive. But in practice this is very common and is known among the people as the degradation problem.

In the case of degradation, we know that shallow networks perform better than deeper samples to which several more layers have been added. So, why not skip these extra layers and at least match them carefully in shallow subnetworks. In this paper, we will do this through a skip connection (identity connection). Also, the use of Batch Normalization ensures that the gradients have healthy norms. Pooling layers are utilized to downsample the input data aiming to reduce the dimensions of the feature maps, resulting reduced height and width.

Now, we are ready to design the X-module as depicted in Fig. 2. To design a X-module, we paid special attention to the memory consumed. The point here is that stacking a large number of layers on top of each other and adding a large number of filters will increase computing, memory consumption, and increase the chances of overfitting. To overcome these problems, we employed the idea of a wider network instead of just a deeper one, increasing the depth and width of the network. The result of this idea was the emergence of the *X module*.

1×1 convolution introduced by Lin et al. [11], which is referred to as "Network in Network". Through a dimensionality reduction with stacked 1×1 convolution filters, X-modules allow more efficient computation and deeper networks. These modules are designed to overcome the problem of vanishing gradients, computational cost, as well as overfitting, among other problems. We put several kernel filters within the CNN, and instead of stacking them in sequential, ordering them to

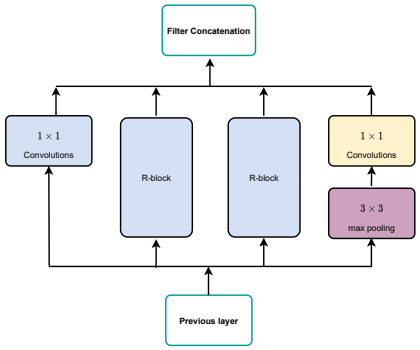


Fig. 2. X-module. This module is used to widen a network.

perform on the same level. 1×1 convolution filters enable the network to learn more information.

The proposed architecture is illustrated in Fig. 3, where X-modules are stacked linearly and the feature map resulted from the last X-module is forwarded to the global average pooling layer. An interesting feature of the proposed architecture is the use of two intermediate classifiers inspired by GoogLeNet. This feature does not exist in other deep architectures for text data. The intuition behind this addition is to prevent the "dying out" of the middle part of the network. The two intermediate classifiers apply the softmax activation function to the outputs of the X-modules, thus computing an intermediate loss on the same labels. The final loss function is the weighted sum of the intermediate losses and the actual loss. The weight used in this paper is 0.4 for each intermediate loss.

IV. Experimental results

Multi-label classification is challenging not only in the development of methods, but also in the evaluation of results. Because it is difficult to distinguish which of the following errors is worse: one sample with three incorrect labels versus three samples each with one incorrect label [12]. Therefore, we must use criteria to examine the performance of our classifier from different angles [12, 14, 15].

Jaccard index- The Jaccard index, also known as multi-labeling accuracy, is computed as the number of correctly predicted labels divided by the total actual predicted labels:

$$Jaccard\ index = \frac{1}{N}\sum_{i=1}^{N}\frac{Y_i \cap \widehat{Y}_i}{Y_i \cup \widehat{Y}_i} \tag{1}$$

where Y is the actual label, $\widehat{Y}_i$ is the predicted label, and N is the total number of samples.

Hamming Loss- Hamming error calculates the symmetric difference ($\triangle$) between the predicted label set *h* and the actual label set *y* for each label *l* and sample *d*. $||$ counts all the elements in this discrepancy (representing erroneous predictions) and normalizes them by the number of labels and samples. In other words, Heming's error can be considered a fraction of labels that have been mispredicted. That is, a fraction of the wrong labels to the total number of labels. Its equation can be written as follows:

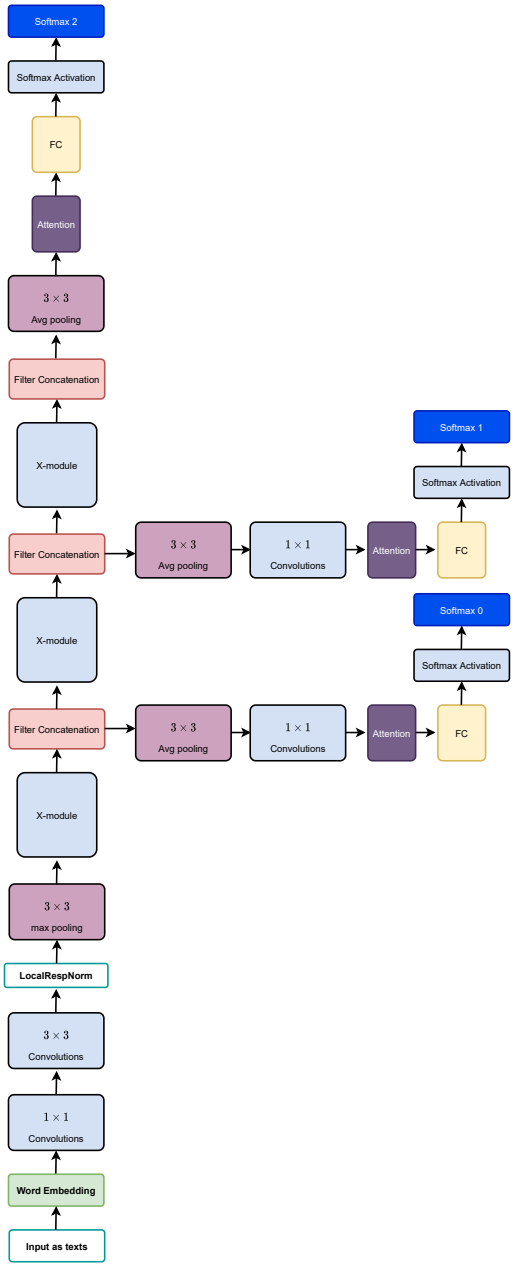


Fig. 3. The proposed architecture

$$Hamming\ Loss\ (h) = \frac{1}{dl}\sum_{i=1}^{d}\sum_{j=1}^{l}|h_{ij} \triangle y_{ij}| \tag{2}$$

Let TP, TN, FP, and FN denote individual classes' true positive, true negative, false positive, and false negative of a model, respectively. The micro-average precision and recall is calculated as follows:

$$Precision\ Micro\ Average = \frac{\sum_{j=1}^{i} TP_j}{\sum_{j=1}^{i} TP_j + \sum_{j=1}^{i} FP_j} \tag{3}$$

$$Recall\ Micro\ Average = \frac{\sum_{j=1}^{i} TP_j}{\sum_{j=1}^{i} TP_j + \sum_{j=1}^{i} FN_j} \tag{4}$$

The macro-average precision and recall is calculated as arithmetic mean of individual classes' precision and recall scores.

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

$$Precision\ Macro\ Averages = \frac{\sum_{m=1}^{k} Precision_k}{k} \tag{7}$$

$$Recall\ Macro\ Averages = \frac{\sum_{m=1}^{k} Recall_k}{k} \tag{8}$$

The micro-macro average F1-score is calculated as harmonic mean of individual classes' F1-score.

$$F1 - score = \frac{2 \times precision \times recall}{precision + recall} \tag{9}$$

Macro F1-Score is the harmonic mean of macro-Precision and macro-Recall; and micro F1-Score is also the harmonic mean of micro-Precision and micro-Recall.

After designing the model, it should now be evaluated on the dataset by the stated criteria. All evaluations are based on 5-fold cross-validation. These evaluations can be seen on the 9-class dataset in Tables I to III. In Table I, the constructed model, called CNN, is applied to the data without any word embedding techniques. Then two different word embedding techniques, *Fast Text* and *Glove*, are applied as input layer on the proposed model. As the results show, in most of the metrics used for evaluation, CNN + GLOVE performs better than the others. In Table II, the proposed models for identifying different emotions are evaluated. As mentioned earlier, there is no way to recognize at least two emotions. Therefore, it is not possible to compare methods. However, Table III shows a comparison of the proposed model with a number of algorithms in this field. As can be seen in the table, the accuracy of the proposed method is better than the compared methods.

V. Conclusion

A system that wants to act like a human being in dealing with emotions and be able to communicate better with human beings needs to recognize different emotions in a text in a multi-labeled way. Because the human brain has the ability to recognize the various emotions that appear in a text, simultaneously. Therefore, having such a system requires a high-performance model that can perform multi-label classification more accurately. In this paper, employing convolutional layers, we present a new model for emotions detection. We equipped the proposed model using the attention layer, and also word embedding techniques and then trained it on the entire multi-label dataset. The experimental results showed good accuracy of the proposed model compared to existing methods.

TABLE I
THE RESULTS ON THE DIFFERENT METRICS

model/metric	Jaccard	Hamming Loss	micro Precision	macro Precision	micro Recall	macro Recall	micro F1	macro F1
CNN	0.57860	0.16528	0.66686	0.49014	0.68244	0.52981	0.67296	0.50280
CNN + FAST TEXT	0.63931	**0.12279**	0.69672	0.50933	**0.79751**	**0.68576**	0.72157	0.55152
CNN + GLOVE	**0.64663**	0.12864	**0.71257**	**0.52368**	0.78746	0.66818	**0.72739**	**0.56238**

TABLE II
THE F1-SCORE FOR DIFFERENT CLASSES

model/class	anger	fear	joy	love	sadness	surprise	thankfulness	disgust	guilt
CNN	0.72902	0.52036	0.72150	0.75194	0.56389	0.46675	0.0	0.73834	0.0
CNN + FAST TEXT	0.78322	0.72270	0.77238	0.83824	0.57292	0.38448	0.15204	0.79034	0.0
CNN + GLOVE	0.80004	0.71673	0.78133	0.83710	0.59937	0.44019	0.15703	0.80991	0.0

TABLE III
COMPARISON RESULTS

Ref.	Dataset	Class	Accuracy	Hamming Loss	micro F1	macro F1
[16]	SemEval-2018	11	0.431	-	0.523	0.413
[14]	BMET	6	0.5856	0.1276	0.6719	-
[14]	SemEval-2018	11	0.5919	0.1211	0.7089	-
[14]	CBET	9	0.5823	0.0739	0.6508	-
[4]	CBET	9	-	-	0.4949	-
[8]	NLPCC2018	5	-	0.1249	0.6076	0.4831
[8]	REN-CECPS	9	-	0.1849	0.6332	0.4923
The proposed CNN + Glove	New dataset (SemEval-2018 + CBET)	9	0.64663	0.12864	0.72739	0.56238

REFERENCES

[1] Park, Ji Ho. "Finding good representations of emotions for text classification." arXiv preprint arXiv:1808.07235 (2018).

[2] Sailunaz, Kashfia, and Reda Alhajj. "Emotion and sentiment analysis from Twitter text." Journal of Computational Science 36 (2019): 101003.

[3] Shaheen, Shadi, Wassim El-Hajj, Hazem Hajj, and Shady Elbassuoni. "Emotion recognition from text based on automatically generated rules." In 2014 IEEE International Conference on Data Mining Workshop, pp. 383-392. IEEE, 2014.

[4] Shahraki, Ameneh Gholipour, and Osmar R. Zaiane. "Lexical and learning-based emotion mining from text." In Proceedings of the international conference on computational linguistics and intelligent text processing, vol. 9, pp. 24-55. 2017.

[5] Jabreel, Mohammed, and Antonio Moreno. "A deep learning-based approach for multi-label emotion classification in tweets." Applied Sciences 9, no. 6 (2019): 1123.

[6] Park, Seo-Hui, Byung-Chull Bae, and Yun-Gyung Cheong. "Emotion recognition from text stories using an emotion embedding model." In 2020 IEEE International Conference on Big Data and Smart Computing (BigComp), pp. 579-583. IEEE, 2020.

[7] Suhasini, Matla, and Badugu Srinivasu. "Emotion detection framework for twitter data using supervised classifiers." In Data Engineering and Communication Technology, pp. 565-576. Springer, Singapore, 2020.

[8] Deng, Jiawen, and Fuji Ren. "Multi-label emotion detection via emotion-specified feature extraction and emotion correlation learning." IEEE Transactions on Affective Computing (2020).

[9] Mohammad, Saif, Felipe Bravo-Marquez, Mohammad Salameh, and Svetlana Kiritchenko. "Semeval-2018 task 1: Affect in tweets." In Proceedings of the 12th international workshop on semantic evaluation, pp. 1-17. 2018.

[10] Almani, Nada M. "Aspect-based sentiment analysis for Arabic reviews using deep learning." PhD diss., University of Surrey, 2019.

[11] Lin, Min, Qiang Chen, and Shuicheng Yan. "Network in network." arXiv preprint arXiv:1312.4400 (2013).

[12] Cabrera-Diego, Luis Adrián, Nik Bessis, and Ioannis Korkontzelos. "Classifying emotions in Stack Overflow and JIRA using a multi-label approach." Knowledge-Based Systems 195 (2020): 105633.

[13] Rodríguez Medina, S., "Multi-Label Text Classification with Transfer Learning for Policy Documents," Master's Thesis, Uppsala University, 2019.

[14] Huang, Chenyang, Amine Trabelsi, Xuebin Qin, Nawshad Farruque, and Osmar R. Zaïane. "Seq2emo for multi-label emotion classification based on latent variable chains transformation." arXiv preprint arXiv:1911.02147 (2019).

[15] Muhammad Hammad Fahim Siddiqui, "Social Media Emotion Detection," Masters Thesis, INSTITUTO POLITÉCNICO NACIONAL, 2020.

[16] Mulki, Hala, Chedi Bechikh Ali, Hatem Haddad, and Ismail Babaoğlu. "Tw-star at semeval-2018 task 1: Preprocessing impact on multi-label emotion classification." In Proceedings of the 12th International Workshop on Semantic Evaluation, pp. 167-171. 2018.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Determining the Severity of Depression in Speech Based on Combination of Acoustic-Space and Score-Space Features

Nasrin Hamiditabar[1], Abdolah Chalechale[2], Seyed Jahanshah Kabudian[3]

[1] *Department of Electrical and Computer Engineering, Razi University, Kermanshah, IRAN,* N.hamiditabar@gmail.com
[2] *Department of Electrical and Computer Engineering, Razi University, Kermanshah, IRAN,* Chalechale@razi.ac.ir
[3] *Department of Electrical and Computer Engineering, Razi University, Kermanshah, IRAN,* Kabudian@razi.ac.ir

***Abstract*—Mental diseases like depression create disorders in vocal limbs and artificial intelligence researchers believe that they can helppsychologists and psychiatrists to measure depression level based on voice features in the patient's speech. The proposed method in thispaper combines acoustic-space and score-space features to estimate Beck's Depression Index (BDI-II) using several regression techniques.To this end, Mel Frequency Cepstral Coefficient (MFCC) features and glottal waveform, pitch, perceptual linear predictive cepstralcoefficients (PLPC), IIR-CQT Frequency Cepstral Coefficients (ICMC), Minimum-Variance Distortionless Response (MVDR), Gaborfilter bank features (GBFB), and separated Gabor filter bank features (SGBFB) along with the scores obtained from each and all of themwas used individually and in combination. The results indicated that combining Long-Term ICMC and Long-Term MFCC features andtheir GMM-based scores led to the best estimate. In addition, based on the AVEC2014 open-access database, the approach adopted hereoutperformed those of many other studies so that a lower mean absolute error (MAE) was obtained.**

***Index Terms* —Speech depression detection, acoustic-space features, score-space features, effects of depression on human's voice.**

I. INTRODUCTION

Clinical depression is a highly serious mental disorder with a growing prevalence. The disease affects speech production through stimulating vocal muscles and vocal folds [1]. As shown by studies, among the features of depressed individuals' voice are reduced verbal output, voice monotony, reduced rate of rapid and significant changes in speed, slower speech, less variation of tone, increased pause duration, long initiation time latency, whispering before uttering a word, occasional stuttering, and decreased variation of pitch are notable [1-3]. There are three methods to diagnose depression including voice processing, video processing, and text processing. Since transcription in text processing and reluctance to disclose one's identity in video processing are serious challenges, there has been a surge in the number of studies on diagnosing depression through analyzing voice signals as a safe and prosperous method.

Several studies have designed an automatic depression detection only based on vocal features such as short-term energy, pitch (F0), intensity, zero-crossing rate, jitter, shimmer, formants, and harmonic-to-noise ratio (HNR) [4-6]. Some of prosodic features can be extracted using a harmonic model to have a more accurate estimate of the severity of depression [7]. To diagnose depression, prosodic vocal tract and glottal features were extracted from signal pieces and different combinations of the features were used for gender-based categorization [8]. The results indicated that, compared to other features, the features based on glottal waveform improved discriminating depressed and healthy speeches. Some diseases, including depression, cause disorder in one's voice and create a wider time diversity in the inflicted individuals' speech. It was shown in [6, 9] that some high-level time-domain features were more successful in determining the disease.

Multi-channel classification systems based on SVM classification have been used to solve depression diagnosing problems along with a weighting parameter for class estimation [10, 4]. In [11, 12], four types of audio features including glottal, prosodic, spectral and teager energy operators were extracted from a set of data and recorded by a multichannel prediction system. The results showed that the features based on TEO-CB-Auto-Env clearly outperformed other features and a combination of other features in diagnosing depression. The ROT features (Ratio of Transient parts) were used in [2] to distinguish depressed and healthy individuals. The low ROT values in depressed individuals mean that they cannot make fast changes in their speech. According to [13], depressed individuals' voice has a lower animation level and this decreases the number of pitches positioned in wider distances and a bit sharper.

In [14], formant analysis was done at vowel level and the results indicated that the mean formant level in female and male voices had contrasting tendencies, which makes this approach a reliable way to diagnose depression. It is believed that short-time mean energy, short-time mean amplitude, increase in first formant, and decrease in second formant are related to depression level [15]. Studies have shown that the eigen value

of correlation matrix and covariance matrix of formant frequency and MFCC can be used to diagnose depression [16, 17]. Voice pitch, energy, root mean squares, and intensity have the best performance on a set of data and yielded the best diagnosis of depression with combined classifications based on GMM and SVM models [18]. Audio features such as intensity, pitch, zero crossing, F0, eight Line-spectral pairs (LSPs), MFCC, and SVM classification were used to diagnose postnatal depression [3]. To diagnose depression, a feature vector for each individual was developed through combining voice, video, and semantic features and after multi-step feature selection, the lowest level of error was achieved with only 46 features for modeling based on SGD algorithm [19]. In [20], low-level descriptors audio features were used to diagnose depression level and by implementing statistical functions, a feature vector was developed for each individual, which was used on a five-layer neural network. After training the network and determining the final weights, the output of the second fully-connected layer was fed into a deep neural network. Logistic regression model and random forest model accompanied with low-level descriptor features (LLDs) can be used as an efficient tool to diagnose depression [21].in [22] Audio and video features and features based on text were used on Gaussian Staircase Regression (GSR) model, which is based on GMM, and the audio system yielded better results in comparison to two other systems.

The method proposed in this paper extracts long-term features out of diverse voice features after implementing a set of preprocessing. In addition, it uses the scores obtained in the classification model (GMM) based on the audio features.

The rest of the paper is organized as follows; section II introduces the proposed method; section III reports the results of experiments using the proposed approach; and finally section IV concludes the paper and also presents some future works.

II. PROPOSED METHOD

The proposed method is designed to efficiently estimate depression level. It starts with preprocessing audio files and then extracting significant features. Afterward, BDI-II is measured using diverse regression techniques out of the subjects' speech to distinguish depressed and healthy individuals. The framework of the proposed depression estimation method is pictured in Figure 1. The following sub-sections give more details of the proposed approach.

A. Preprocessing

Before extracting the features proportional to the signal adopted, a few preprocessing stages are needed. After examining Freeform audio files using a proper audio editor software, we found that five audio files in the training dataset, three audio files in the development data, and one audio file in test data were corrupted, which were removed from the study. Afterward, the whole data was down-sampled to 16 kHz and the silence regions were spotted and truncated using VAD algorithm. To deal with probable issues in recording the audio files, all the files were normalized before further analyses.

B. Feature extraction

Feature extraction has a highly important role in depression analysis systems. The knowledge gained through several decades of studies on speech processing can help us. To create a strong speech depression detection system, long-term acoustic features and Gaussian Mixture Model (GMM)-based scores based on speech signal and speech signal waveform-based features such as MFCC, glottal waveform features, pitch, PLPC [23], ICMC [24], MVDR [25], GBFB and SGBGB [26] were used. After extracting the short-term features, long-term features (mean, variance, skewness, and kurtosis) that contained the key information of feature matrix and at the same time decreased the volume and complicacy of computations were used. Then, to extract the score, the audio files of the training set were categorized into depressed and healthy groups. Depressed individuals were defined as the cases with DBI-II scores $\geq$14 and healthy individuals were defined as the cases with BDI-II scores <14. The depressed group was marked with classification label 1 and the healthy group was mark by classification label 0. Afterward, the scores obtained from depressed and healthy GMM models trained on short-term feature matrix of both groups were used as a 2-dimensional feature vector for each audio file. The final feature fed into regression stage was a combination of long-term acoustic feature vector and GMM-based score vector.

C. *Regression*

Following general mean and variance normalization (z-score normalization), the extracted features were modeled using diverse regression methods to estimate BDI-II assessment scale. Regression methods like Support Vector Regression (SVR) [27] with gaussian, linear, and polynomial kernels, Ridge regression [28], Huber [29], Tweedie [30], Gradient Boosting [31], k-NN, Decision Tree, Bayesian, and Random Forest were employed.

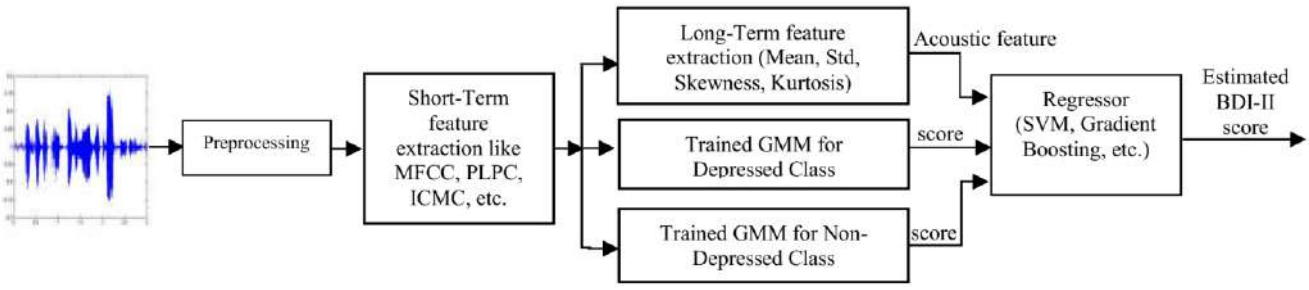


Figure 1. The Block diagram of the proposed depression estimation system

III. EXPERIMENTAL RESULTS

A. AVEC2014 Dataset

Here, AVEC2014 (Audio/Visual Emotion Challenge 2014) [32], a subset of AVEC2013 German language database was used. The audio files of 84 individuals with mean age of 31.15±12.3 (18-63) years were used. Each record contained two tasks of reading a part of "North Wind and Sun" story and a Freeform speech of answering a virtual interviewer. Duration of the audio files was in 6s to 4min range. Each record was tagged using BDI-II self-assessment questionnaire. The audio files in AVEC2014 were categorized into three parts of training, development, and test each with 100 records for both tasks. Since Freeform speech is a better way than reading to examine mental and spiritual behaviors of individuals, the experiments were done using the Freeform speech data. To train the system, training and development data were combined, and to assess the system, the test data were used.

B. Evaluation measures

Since the mean absolute error (MAE) is more robust than the Mean Squared Error (MSE) on the outlier data , it was used as the evaluation measure of AVEC2014 challenge, a distance measure between the predicted and ground-truth BDI-II scores, which was computed as follows:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - \bar{y}_i\right| \qquad (1)$$

Where N is the number of samples, y_i is actual value, and $\bar{y}_i$ is the estimated value of the BDI-II for *i* th sample.

C. Long-Term features and scores

To remove the remaining negative slope in the speech signal spectrum, a pre-emphasis filter, which is a first-order filter was implemented on the audio files and to extract 20 MFCC coefficients, 24 Mel bank filters were used. In addition, 20 Mel bank filters and 96 as quality factor were used to extract 20 ICMC coefficients. By implementing Gabor filter bank on Mel-scaled log-spectrogram, 455 GBFB features and 1020 SGBFB features were extracted from the audio signals. In addition, 12 PLPC coefficients, one pitch feature, 36 MVDR features, and 68 glottal features were used in the study. As to the extracted features, a Hanning window with 32ms frame size and 10ms frame step was used and then the first and second derivatives were added to the features. The GMM model was developed with four gaussians and regularization coefficient of 0.1. To determine the best features, the features performance was recorded for both alone and combined features and out of them, the regression methods with the best results are listed in TABLE I. As listed in TABLE II, the combination of features and scores improved the assessment index and as listed in TABLE III, CMN-CVN normalization on the features matrix led to a notable improvement. The best results are bolded in the tables.

D. Performance comparison

A comparison of the results with those of some other studies is listed in TABLE IV. The proposed system in this study had a better estimate of BDI-II scale so that the assessment measure was improved.

IV. CONCLUSION AND FUTURE WORKS

In this paper, a speech depression detection approach was proposed that uses both acoustic-space and score-space features to measure the severity of depression. Through examining the features based on speech signal and speech glottal waveform (e.g. MFCC, Glottal, PITCH, PLPC, ICMC, MVDR, GBFB, SGBFB) and the score of GMM model obtained from each one of the features and a combination of them, it was found that the long-term ICMC and long-term MFCC features and the GMM scores yielded the best estimate. The MAE evaluation measure was also decreased to 7.63. In addition, given the high noise level in some of the audio files, noise reduction and normalization had a positive effect on depression severity diagnosis method. In the future, the authors wish to employ deep learning techniques to improve diagnosing power of the proposed approach.

TABLE I. MAE RESULTING FROM THE LONG-TERM FEATURES (SINGLE AND COMBINED) USING DIFFERENT REGRESSION METHODS

features	Random forest	Gradient Boosting	Tweedie	Bayesian	Ridge	SVR (RBF)
PITCH	9.571	10.897	9.378	9.459	9.466	8.302
GLOTTAL	9.068	9.318	8.930	8.499	8.236	9.412
MFCC	9.190	10.602	8.519	8.715	8.604	9.385
PLPC	8.955	9.174	8.851	8.824	8.191	8.428
MVDR	9.386	10.840	9.590	8.934	9.002	8.970
ICMC	8.617	10.137	8.533	8.426	8.425	9.191
GBFB	8.753	9.822	9.۴۳۰	8.788	9.255	9.201
SGBFB	8.566	10.049	9.162	8.683	9.102	9.128
MFCC PLPC PITCH	9.077	8.168	8.705	**8.143**	8.721	8.926
MFCC PLPC	9.152	8.584	8.726	8.166	8.743	8.939
PITCH PLPC	8.823	9.412	8.968	8.826	9.083	8.491
GLOTTAL PLPC MFCC PITCH	8.944	9.449	9.035	8.339	8.464	9.281
ICMC MFCC	8.836	10.083	8.741	8.345	8.777	9.251
ICMC SGBFB MFCC PLPC MVDR	8.531	9.365	9.020	8.622	8.707	9.101
MFCC PLPC GLOTTAL ICMC	8.516	8.349	8.791	8.254	8.351	9.244

TABLE II. MAE RESULTING FROM THE COMBINATION OF LONG-TERM FEATURES AND SCORES, WITHOUT NORMALIZATION

features	Random forest	Gradient Boosting	Tweedie	Bayesian	Ridge	SVR (RBF)
MFCC PLPC PITCH	8.659	9.216	8.997	9.383	8.718	**8.056**
MFCC PLPC	8.701	9.245	9.077	9.098	8.713	9.151
PITCH PLPC	8.769	9.591	9.173	9.388	9.029	8.756
PITCH PLPC MFCC GLOTTAL	8.695	10.386	9.222	9.307	8.533	8.470
ICMC MFCC	8.770	8.273	9.031	9.002	8.745	9.203
ICMC SGBFB MFCC PLPC MVDR	8.696	10.418	10.822	9.300	8.902	9.480
MFCC PLPC GLOTTAL ICMC	8.498	10.013	10.400	9.466	8.444	9.769

TABLE III. MAE RESULTING FROM THE COMBINATION OF LONG-TERM FEATURES AND SCORES, AFTER NORMALIZATION

features	Random forest	Gradient Boosting	Tweedie	Bayesian	Ridge	SVR (RBF)
MFCC PLPC PITCH	8.731	9.880	8.775	8.187	8.545	8.896
MFCC PLPC	8.863	9.530	8.744	8.180	8.562	8.825
PITCH PLPC	9.106	9.640	8.992	8.771	8.917	8.976
PITCH PLPC MFCC GLOTTAL	8.635	10.184	8.845	8.386	8.314	9.310
ICMC MFCC	8.622	**٧.٦٣٠**	8.792	8.410	8.614	8.938
ICMC SGBFB MFCC PLPC MVDR	8.519	10.257	9.041	9.087	8.726	9.078
MFCC PLPC GLOTTAL ICMC	8.456	9.338	8.860	8.311	8.196	9.273

TABLE IV. COMPARISON OF THE RESULTS OF THE PROPOSED METHOD WITH SOME OTHER WORKS

	Title	Year	MAE
1	AVEC 2014 – 3D dimensional affect and depression recognition challenge[32]	2014	10.04
2	Model fusion for multimodal depression classification and level detection [33]	٢٠١٤	9.82
3	Fusing affective dimensions and audio-visual features from segmented video for depression recognition [34]	٢٠١٤	9.35
4	Automatic depression scale prediction using facial expression dynamics and regression [35]	٢٠١٤	9.10
5	The SRI AVEC-2014 evaluation system [36]	2014	8.83
6	Influence of speaker de-identification in depression detection [37]	٢٠١٧	8.08
7	Speech vs. text: a comparative analysis of features for depression detection systems [38]	٢٠١٦	8.59
8	Depression estimation using audio visual features and fisher vector encoding [39]	٢٠١٤	8.40
9	Automated depression analysis using convolutional neural networks from speech [40]	2018	8.19
10	Ensemble CCA for continuous emotion prediction [41]	2014	٧,٩٦
11	Hybrid network feature extraction for depression assessment from speech [42]	٢٠٢٠	7.94
12	Multimodal spatiotemporal representation for automatic depression level detection [43]	٢٠٢٠	7.65
13	*Proposed Method*		٧,٦٣

REFERENCES

[1] N. Cummins, et al, "A review of depression and suicide risk assessment using speech analysis," Speech Communication, vol. 71, July 2015, pp. 10-49.

[2] G. Kiss, A. B. Takács, D. Sztahó, K. Vicsi, "Detection possibilities of depression and parkinson's disease based on the ratio of transient parts of the speech, " 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom 2018), 2018, pp.165-168.

[3] J. Wang, et al, "Detecting postpartum depression in depressed people by speech features," Springer International Conference on Human Centered Computing(HCC), vol. 10745, 2018, pp. 433-442.

[4] H. Long, et al, "Detecting depression in speech comparison and combination between different speech types," IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2017, pp. 1052-1058.

[5] Y. Yang, C. Fairbairn, and J. F. Cohn, "Detecting depression severity from vocal prosody," IEEE Transactions On Affective Computing, vol. 4, 2013, pp. 142 – 150.

[6] A. Roniotis and M. Tsiknakis, "Detecting depression using voice signal extracted by chatbots a feasibility study," ICST Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, 2018, pp. 386–392.

[7] M. Asgari, I. Shafran and L. B. Sheeber, "Inferring clinical depression from speech and spoken utterances," IEEE International Workshop on Machine Learning for Signal Processing (MLSP), 2014.

[8] E. Moore, M. A. Clements, J. W. Peifer and L. Weisser, "Critical analysis of the impact of glottal features in the classification of clinical depression in speech," IEEE Transactions On Biomedical Engineering, vol. 55, 2008, pp. 96 – 107.

[9] M. C. Astorgano, D. E. Mancebo and C. G. Ferreras, "Acoustic characterization and perceptual analysis of the relative importance of prosody in speech of people with down syndrome," Speech Communication, vol. 99, 2018, pp. 90-100.

[10] H. Jiang, et al, "Investigation of different speech types and emotions for detecting depression using different classifiers," Speech Communication, vol. 90, 2017, pp. 39-46.

[11] L. S. A. Low, N. C. Maddage, M. Lech, L. B. Sheeber, and N. B. Allen, "Detection of clinical depression in adolescents' Speech During Family Interactions," IEEE Transactions On Biomedical Engineering, vol. 58, 2011, pp. 574 - 586.

[12] K. E. B. Ooi, M. Lech and N. B. Allen, "Multichannel weighted Speech classification system for prediction of major depression in adolescents," IEEE Transactions On Biomedical Engineering, vol. 60, 2013, pp. 497 - 506.

[13] V. Mitra and E. Shriberg, "Effects of feature type, learning algorithm and speaking style for depression detection from speech," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2015, pp. 4774-4778.

[14] N. Cummins (B), B. Vlasenko, H. Sagha and B. Schuller, "Enhancing speech-based depression detection through gender dependent vowel-level formant features," Springer International Publishing Conference on Artificial Intelligence in Medicine, 2017, pp. 209–214.

[15] J. Zhao, W. Su , J. Jia , C. Zhang and T. Lu, " Research on depression detection algorithm combine acoustic rhythm with sparse face recognition," Springer Science+Business Media, LLC, part of Springer Nature 2017, 2017, pp. 7873–7884.

[16] J. R. Williamson et al, "Vocal biomarkers of depression based on motor incoordination," 3rd ACM International Workshop on Audio/visual emotion challenge, October 2013, pp.41-48.

[17] J. R. Williamson, T. F. Quatieri, B. S. Helfer, G. Ciccarelli and D. D. Mehta, "Vocal and facial biomarkers of depression based on motor incoordination and timing," the 4th International Workshop on Audio/Visual Emotion Challenge, November 2014, pp. 65–72.

[18] Sh. Alghowinem, et al, "A comparative study of different classifiers for detecting depression from spontaneous speech," IEEE International Conference on Acoustics, Speech and Signal Processing, 2013, pp. 8022-8026.

[19] Y. Gong and C. Poellabauer, "Topic modeling based multi-modal depression detection," 7th Annual Workshop on Audio/Visual Emotion Challenge, October 2017, pp. 69-76.

[20] L. Yang, et al, "Hybrid depression classification and estimation from audio video and text information," 7th Annual Workshop on Audio/Visual Emotion Challenge , October 2017, pp. 45-51.

[21] L. Zhang, J. Driscol, X. Chen and R. H. Ghomi, "Evaluating acoustic and linguistic features of detecting depression sub-challenge dataset," 9th International on Audio/Visual Emotion Challenge and Workshop, October 2019, pp. 47-53.

[22] T. Dang, et al, "Investigating word affect features and fusion of probabilistic predictions incorporating uncertainty in AVEC 2017," 7th Annual Workshop on Audio/Visual Emotion Challenge , October 2017, pp. 27-35.

[23] H. Hermansky, "Perceptual linear predictive (PLPC) analysis of speech," The Journal of the Acoustical Society of America, vol. 87, 1990, pp. 1738-1752.

[24] D. hector, et al, "Further optimisations of constant Q cepstral processing for integrated utterance and text-dependent speaker verification," IEEE Spoken Language Technology Workshop (SLT), 2016, pp. 179-185.

[25] U. H. Yapanel and J. H. L. Hansen, "A new perceptually motivated MVDR-based acoustic front-end (PMVDR) for robust automatic speech recognition," Speech Communication, vol. 50, 2008, pp. 142-152.

[26] M. R. Schädler and B. Kollmeier, "Separable spectro-temporal gabor filter bank features: reducing the complexity of robust features for automatic speech recognition," The Journal of the Acoustical Society of America, vol. 137, 2015, pp. 2047–2059.

[27] M. Awad and R. Khanna, "Support vector regression," Efficient Learning Machines, 2015, pp. 67-80.

[28] D. Hilt, "Ridge, A computer program for calculating ridge regression estimates," Research Note NE-236. Upper Darby, PA: U.S. Department of Agriculture, Forest Service, Northeastern Forest Experiment Station. 1977.

[29] P. J. Huber, "Robust estimation of a location parameter," Ann. Math. Statis, vol. 35 , March1964, pp. 73-101.

[30] C. C. Kokonendji, W. H. Bonat and R. Abid, "Tweedie regression models and Its geometric sums for (semi-)continuous data, " Wires Computational Statistics, vol. 13, 2019.

[31] Jerome H, "Greedy function approximation: a gradient boosting machine," The Annals of Statistics Published By Institute of Mathematical Statistics , vol. 29, 2001, pp. 1189-1232.

[32] M. Valstar, "AVEC 2014 – 3D dimensional affect and depression recognition challenge," 4th International Workshop on Audio/Visual Emotion Challenge, November 2014, pp. 3-10.

[33] M. Senoussaoui, M. Sarria-Paja, J. F. Santos and T. H. Falk, "Model fusion for multimodal depression classification and level detection," 4th International Workshop on Audio/Visual Emotion Challenge, November 2014, pp. 57-63.

[34] H. Pérez, et al, "Fusing affective dimensions and audio-visual features from segmented video for depression recognition," 4th International Workshop on Audio/Visual Emotion Challenge, November 2014, pp. 49-55.

[35] A. Jan, "Automatic depression scale prediction using facial expression dynamics and regression," 4th International Workshop on Audio/Visual Emotion Challenge, November 2014, PP. 73-80.

[36] V. Mitra, "The SRI AVEC-2014 evaluation system," 4th International Workshop on Audio/Visual Emotion Challenge , November 2014, pp. 93-101.

[37] P. Lopez-Otero, "Influence of speaker de-identification in depression detection," Signal Processing,vol. 11, 2017, pp. 1023-1030.

[38] M. R. Morales, "Speech vs. text: a comparative analysis of features for depression detection systems," IEEE Spoken Language Technology Workshop (SLT), 2016 , pp. 136-143.

[39] V. Jain, "Depression estimation using audiovisual features and fisher vector encoding," 4th International Workshop on Audio/Visual Emotion Challenge, November 2014, pp. 87-91.

[40] L. He and C. Cao, "Automated depression analysis using convolutional neural networks from speech," Journal of Biomedical Informatics, vol. 83, 2017, pp. 103-111.

[41] H. Kaya, F. Çilli and A. A. Salah, "Ensemble CCA for continuous emotion prediction," 4th International Workshop on Audio/Visual Emotion Challenge , November 2014, pp. 19-26.

[42] Z. Zhao, et al, "Hybrid network feature extraction for depression assessment from speech," Interspeech, 2020, pp. 4956-4960.

[43] M. Niu, et al, "Multimodal spatiotemporal representation for automatic depression level detection," IEEE Transactions On Affective Computing, 2020, pp. 1-14.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Domination integrity in intuitionistic fuzzy graphs

1st S. H. Sadati, 2nd A. A. Talebi and 3rd H. Rashmanlou

1st Department of Mathematics, University of Mazandaran, Babolsar, Iran, sadati.asrami.hs@gmail.com
2nd Department of Mathematics, University of Mazandaran, Babolsar, Iran, a.talebi@umz.ac.ir
3rd Department of Mathematics, University of Mazandaran, Babolsar, Iran, rashmanlou.1987@gmail.com

***Abstract*— Among all other graph parameters, graph integrity is a well-known measure of the vulnerability of the network. The effectiveness of the network decreases with the breaking down of some vertices or links. Therefore, a less vulnerable communication network is required for greater stability. Vulnerability is the measure of the resistance of the network after the failure of communication links. A new vulnerability parameter, domination integrity of fuzzy graph is defined to study the stability and domination property of fuzzy graphs. In this paper, we studied the concept of domination integrity in IFGs and investigated some of its properties. We also introduced domination weak integrity in the IFG and studied some of the results.**

***Index Terms*—Intuitionistic fuzz graph, dominating set, domination integrity, domination weak integrity**

I. Introduction

Graph theory plays an important role in applied mathematics and fields such as computer science, engineering, medicine, social sciences, and so on. In this theory, the concept of domination is one of the most widely used concepts used by graphs in the real world. This concept was introduced by Ore [10]. We can find a set of vertices that dominates the remaining vertices in the vertex set, where these vertices are removed from the graph; it may be either connected or disconnected. The dominance parameter leaves the subgraph with the vertices of the set removed. We need another parameter relating to the vertices set and subgraph after deleting. This concept is the integrity of the graph. Integrity accounts for the cordiality of the removed vertices set and connected component after removing that vertices set.

The integrity of graphs was introduced by Barefoot, Entringer, and Swart [4] as a measure of the vulnerability of graphs. This parameter is used to measure the stability of a graph, the number of vertices not in link, and the size of the largest sub network is still in connection. A survey of integrity conducted by Bagga et al. [3]. Domination integrity was defined by Sundareswaran et al. in [25]. They developed the domination integrity concepts in middle graphs [26], in trees [27], in powers of cycles [28], and in gear graphs [29]. The global domination number was introduced by Sampathkumar in [19], and the global domination integrity was introduced and further discussed by Sultan Senan Mahde et al. in [7]. Sibabrata Ray et al. [23] introduced the concept of integrity in weighted graphs. They discussed the NP-Completeness of integrity of weighted graphs.

Rosenfeld [18] introduced the concept of fuzzy graphs from fuzzy relations, defined by Zadeh in 1965 [34]. Various fuzzy graphs were then introduced by researchers. Atanassov [1] introduced the concept of Intuitionistic Fuzzy (IF) relations and Intuitionistic Fuzzy Graphs (IFGs). Some features of intuitionistic fuzzy graphs were introduced by Nagoor gani et al. [8, 9]. Parvathi and Karunambigai [12] gave a definition for IFG as a special case of IFGs defined by Atanassov and Shannon [22]. Talebi et al. [30–33] studied several concepts on fuzzy graphs and interval valued intuitionistic fuzzy graphs. Rashmanlou et al. [13–17] investigated novel concepts in fuzzy graphs.

The concept of domination in the fuzzy graphs was investigated by A. Somasundaram and S. Somoasundaram [24]. Natarajan and Ayyasamy introduced strong (weak) domination in fuzzy graphs [2]. Jayalakshmi and Harinarayanan introduced the Total strong (weak) domination in Fuzzy graphs [5]. Parvathi et al. [11] introduced domination on IFGs. The fuzzy vertex integrity concept was introduced and the vertex integrity value of some standard graphs was discussed. Edge integrity of fuzzy graphs was introduced by Saravanan et al. [20]. Also, Vertex integrity and edge integrity values of join, union, and Cartesian product of fuzzy graphs are discussed by them [21].

In this paper, we studied the concept of domination integrity in IFGs and investigated some of its properties. We also introduced domination weak integrity in the IFG and studied some of the results.

II. Preliinaries

In this section we have an overview of the concepts we need in this article.

Let $G = (V, E)$ be a finite graph without loops and multiple edges. The set of vertices and edges of G are denoted by $V(G)$ and $E(G)$, respectively. The integrity $I(G)$ of a graph G is defined as $I(G) = \min\{|S| + m(G - S) \mid S \subseteq V(G)\}$,

where $m(G-S)$ denotes the order of a maximum component of $G-S$. A subset S of $V(G)$ is said to be an $I-$set if $I(G) = |S| + m(G-S)$. Moreover the pure edge integrity $I_p(G)$ of a graph G is defined as $I_p(G) = \min\{|S| + m_e(G-S) \mid S \subseteq E(G)\}$, where $m_e(G-S)$ denotes the number of edges in a largest component of $G-S$. Also the weak integrity $I_w(G)$ is defined as $I_w(G) = \min\{|S| + m_e(G-S) \mid S \subseteq V(G)\}$.
A subset S of $V(G)$ is called dominating set if for every $v \in V-S$, there exist a $u \in S$ such that v is adjacent to u. The minimum cardinality of a minial dominating set in G is called the domination number of G denoted as $\gamma(G)$ and the corresponding minimal dominating set is called a $\gamma-$set of G.
The domination integrity of a graph is defined as $DI(G) = \min\{|S| + m(G-S)\}$, where S is a dominating set of G and $m(G-S)$ denotes the order of the largest component in $G-S$ and is denoted by $DI(G)$.

Definition II.1. *[34] Let V be nonempty set. The fuzzy set A of V is defined as $A = \{\mu_A(x) \mid x \in V\}$, where $\mu_A : V \to [0,1]$ is the membership function of the fuzzy set.*

Definition II.2. *[1] Let V be a nonempty set. An intuitionistic fuzzy set (IFS) A in V is defined as*

$$A = \{\langle \mu_A(x), \nu_A(x)\rangle \mid x \in V\}$$

where the functions $\mu_A, \nu_A : V \to [0,1]$ define respectively, the degree of membership and degree of nonmembership of the element $x \in V$ such that

$$0 \le \mu_A(x) + \nu_A(x) \le 1, \quad \textit{for all} \quad x \in V$$

The intuitionistic fuzzy relation B on the set $V \times V$ is an intuitionistic fuzzy set (IFS) of the form

$$B = \{\langle \mu_B(x,y), \nu_B(x,y) \mid xy \in V \times V\rangle\}$$

where $\mu_B, \nu_B : V \times V \to [0,1]$ so that

$$0 \le \mu_B(xy) + \nu_B(xy) \le 1, \quad \textit{for all} \quad xy \in V \times V$$

Definition II.3. *[22] An intuitionistic fuzzy graph (IFG) is a pair $\mathcal{G} = (A,B)$, where A is an IFS on the set of vertices V and B is an intuitionistic fuzzy relation on the set of edges $E \subseteq V \times V$ so that*

$$\begin{aligned} \mu_B(xy) &\le \mu_A(x) \wedge \mu_A(y) \\ \nu_B(xy) &\le \nu_A(x) \vee \nu_A(y) \end{aligned}$$

Definition II.4. *[11] Let $\mathcal{G} = (A,B)$ be an IFG. then, the vertex cardinality of V defined by*

$$P = |V| = \sum_{x \in V} \frac{1 + \mu_A(x) - \nu_A(x)}{2}, \quad \textit{for all} \quad x \in V.$$

The vertex cardinality of V is called the order of $\mathcal{G}$.

Definition II.5. *[6] Let $\mathcal{G} = (A,B)$ be an IFG over V. A path of length k in an IFG is a sequence of distinct vertices $x_1, x_2, \cdots, x_k$, such that either one of the following conditions is satisfied:*

1) $\mu_B(x_i x_j) > 0, \quad \nu_B(x_i x_j) > 0$
2) $\mu_B(x_i x_j) = 0, \quad \nu_B(x_i x_j) > 0$
3) $\mu_B(x_i x_j) > 0, \quad \nu_B(x_i x_j) = 0$, *for some i and j.*

The strength of a path is defined as

$$\left\langle \bigwedge_{i=1}^{k} \mu_B(x_{i-1}x_i), \bigvee_{i=1}^{k} \nu_B(x_{i-1}x_i) \right\rangle, \quad i = 1, 2, \cdots, k.$$

The strength of connectedness between x_i and x_j is denoted b $\langle \mu_B^\infty(x_i x_j), \nu_B^\infty(x_i x_j)\rangle$ where

$$\begin{aligned} \mu_B^\infty(x_i x_j) &= \sup\{\mu_B^\infty(x_i x_j) \mid k = 1, 2, \cdots, n\} \\ \nu_B^\infty(x_i x_j) &= \inf\{\nu_B^\infty(x_i x_j) \mid k = 1, 2, \cdots, n\} \end{aligned}$$

Definition II.6. *[11] An edge xy is said to be a strong edge, if $\mu_B(xy) \ge \mu_B^\infty(xy)$ and $\nu_B(xy) \ge \nu_B^\infty(xy)$.*

Definition II.7. *[11] Let $\mathcal{G} = (\mu, \nu)$ be an IFG on V and $x, y \in V$. We say that x dominates y in $\mathcal{G}$ if there exists a strong edge between them. A subset S of V is called a dominating set (DS) in $\mathcal{G}$ if for every $y \in V - S$, there exists $x \in S$ such that x dominates y. S is said to be minial-DS if no proper subset of S is a DS.*
Minimum cardinality among all minimal-DS is called lower domination number of $\mathcal{G}$, and is denoted by $d(\mathcal{G})$.
Maximum cardinality among all minial-DS is called upper domination number of $\mathcal{G}$, and is denoted by $D(\mathcal{G})$.

Definition II.8. *[6] An IFG $\mathcal{H} = (A', B')$ is said to be subgraph of $\mathcal{G} = (A,B)$ if $A' \subseteq A$ and $B' \subseteq B$.*

Definition II.9. *[12] An IFG $\mathcal{G} = (A,B)$ is said to be complete-IFG if*

$$\begin{aligned} \mu_B(xy) &= \mu_A(x) \wedge \mu_A(y) \\ \nu_B(xy) &= \nu_A(x) \vee \nu_A(y), \quad \textit{for every } x, y \in V. \end{aligned}$$

Definition II.10. *[12] The complement of an IFG, $\mathcal{G} = (A,B)$ is an IFG, $\overline{\mathcal{G}} = (\overline{A}, \overline{B})$, where*

i) $\overline{A} = A$,
ii) $\overline{\mu_A}(x) = \mu_A(x)$ *and* $\overline{\nu_A}(x) = \nu_A(x)$, *for all $x \in V$,*
iii) $\overline{\mu_B}(xy) = \mu_A(x) \wedge \mu_A(y) - \mu_B(xy)$ *and* $\overline{\nu_B}(xy) = \nu_A(x) \vee \nu_A(y) - \nu_B(xy)$, *for all $x, y \in V$.*

Definition II.11. *[11] An IFG, $\mathcal{G} = (A,B)$ is sais to complete bipartite if the vertex set V can be partitioned into two non empty sets V_1 and V_2 such that*

i) $\mu_B(xy) = 0$ *and* $\nu_B(xy) = 0$, *if $x, y \in V_1$ or $x, y \in V_2$,*
ii) $\mu_B(xy) > 0$, $\nu_B(xy) > 0$, *or* $\mu_B(xy) = 0$, $\nu_B(xy) > 0$, *or* $\mu_B(xy) > 0$, $\nu_B(xy) = 0$, *if $x \in V_1$ and $y \in V_2$,*
iii) $\mu_B(xy) = \mu_A(x) \wedge \mu_A(y)$ *and* $\nu_B(xy) = \nu_A(x) \vee \nu_A(y)$, *for all $x \in V_1$ and $y \in V_2$.*

III. Domination integrity of intuitionistic fuzzy graphs

In this section, domination integrity in IFG is discussed and its properties are examined.

Definition III.1. *Let $\mathcal{G} = (A, B)$ be a IFG. The domination integrity number of $\mathcal{G}$, denoted by $\widetilde{DI}(\mathcal{G})$, is defined as*

$$\widetilde{DI}(\mathcal{G}) = \min\{|S| + m(\mathcal{G} - S) \mid S \text{ is the } DS \text{ of } \mathcal{G}\} \quad (1)$$

where $|S|$ denoted the vertex cardinality of S and $m(\mathcal{G} - S)$ denotes the maximum order of the components of $\mathcal{G} - S$.
The DS S, which satisfies in (1)*, is called as minimal domination integrity set, and is denoted by $\widetilde{DI}-$set.*

Example III.2. *Consider an IFG $\mathcal{G} = (A, B)$ as shown in "Fig. 1".*
In this graph, the edges ab, bc, cd, bd are strang. The min-

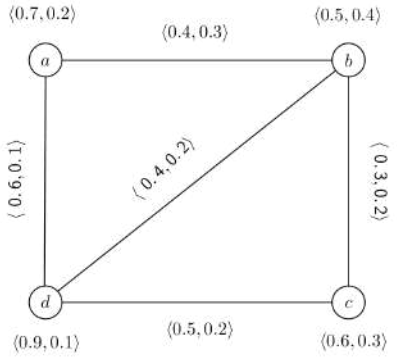


Fig. 1. An IFG

imal DSs are $S_1 = \{b\}$, $S_2 = \{a, c\}$, $S_3 = \{a, d\}$. The domination number is $d(\mathcal{G}) = \min\{|S_1|, |S_2|, |S_3|\} = \min\{0.55, 1.4, 1.65\} = 0.55$. Therefore S_1 is a d-set. On the other hand, we have

$$\widetilde{DI}(\mathcal{G}) = \min\{0.55 + 2.3, 1.4 + 1.45, 1.65 + 1.2\} = 2.85.$$

Thus, S_1, S_2, S_3 are the minimal domination integrity set, but S_1 is the minimum dominatien integrity set in $\mathcal{G}$.

Theorem III.3. *If $\mathcal{G} = (A, B)$ be an IFG, then*

$$d(\mathcal{G}) \leq \widetilde{DI}(\mathcal{G}).$$

Proof: The domination number of IFG depends only on the cardinality of the DS. But the domination integrity number depends upon the DS S and the corresponding maximum order of the component of $\mathcal{G} - S$. This implies $d(\mathcal{G}) < \widetilde{DI}(\mathcal{G})$. Equality holds for an IFG with isolated vertices only for an IFG with only isolated vertices, the whole vertex set is the only DS. Since $m(\mathcal{G} - S) = 0$ implies $d(\mathcal{G}) = \widetilde{DI}(\mathcal{G})$. Hence, $d(\mathcal{G}) \leq \widetilde{DI}(\mathcal{G})$. ∎

Corolary III.4. *The domination integrity parameter is a monotonically increasing parameter i.e. If $\mathcal{H}$ be a subgraph of an IFG $\mathcal{G} = (A, B)$, then*

$$\widetilde{DI}(\mathcal{H}) \leq \widetilde{DI}(\mathcal{G}).$$

Remark III.5. *For any IFG $\mathcal{G} = (A, B)$,*

$$d(\mathcal{G}) \leq \widetilde{DI}(\mathcal{G}) \leq p.$$

Remark III.6. *For any IFG $\mathcal{G} = (A, B)$,*

$$\widetilde{DI}(\mathcal{G}) + \widetilde{DI}(\widetilde{\mathcal{G}}) \leq 2p.$$

The inequality is trivial and equality holds for complete IFG and its complement.

Theorem III.7. *Let $\mathcal{G} = (A, B)$ be an IFG. $\widetilde{DI}(\mathcal{G}) = p$ if and only if $\mathcal{G}$ is either complete-IFG or complement of complete-IFG.*

Proof: Let $\mathcal{G} = (A, B)$ be a complete-IFG. Then, any subset of V is DS of $\mathcal{G}$. Eliminating any subset of V the remaining graph is a single component consisting of all the remaining vertices. Thus, $\widetilde{DI}(\mathcal{G}) = p$ suppose $\mathcal{G}$ is complement of the complete-IFG. Then, $\mathcal{G}$ is an IFG with set of isolated vertices. The whole vertex set is the only DS of $\mathcal{G}$. since $m(\mathcal{G} - S) = 0$, then domination integrity of $\mathcal{G}$ is p. Therefore,

$$d(\mathcal{G}) = \widetilde{DI}(\mathcal{G}) = p.$$

Conversely, let $\widetilde{DI}(\mathcal{G}) = p$ and S be a DS of $\mathcal{G}$. If $\mathcal{G} - S$ contains more than one component, then sum of cardinality of S and $m(\mathcal{G} - S)$ must be less than p. So, $\mathcal{G} - S$ must be a single connected component. This is hold for all DS, in particular for any singleton vertex set. Thus, this singleton vertex set dominates all the remaining vertices of $\mathcal{G}$. Thus $\mathcal{G}$ is a complete-IFG.

∎

Theorem III.8. *Let $\mathcal{G} = (A, B)$ be a complete bipartite-IFG with the bipartion V_1 and V_2. The domination integrity of $\mathcal{G}$ is*

$$\widetilde{DI}(\mathcal{G}) = \min\left\{|V_1| + \bigvee_{x \in V_2} A(x), \quad |V_2| + \bigvee_{x \in V_1} A(x)\right\}$$

Proof: Let $\mathcal{G}$ be a complete bipartite-IFG with the bipartion V_1 and V_2. We know, the domination set S of $\mathcal{G}$ is either V_1 or V_2 or one vertex from V_1 and V_2. If S is V_1, then $\mathcal{G} - S$ is a collection of isolated vertices of V_2. Therefore, $m(\mathcal{G} - S) = \bigvee_{x \in V_2} A(x)$. Similarly, if S is V_2, then $\mathcal{G} - S$ is a collection of isolated vertices of V_1 which implies $m(\mathcal{G} - S) = \bigvee_{x \in V_1} A(x)$. If S has one vertices from V_1 and other from V_2, then $\mathcal{G} - S$ remains a single connected component. Considering the minimum value for the above three DSs, we may ignore the DS S, having one vertex in V_1 and another one from V_2. Thus,

$$\widetilde{DI}(\mathcal{G}) = \min\left\{|V_1| + \bigvee_{x \in V_2} A(x), \quad |V_2| + \bigvee_{x \in V_1} A(x)\right\}.$$

Theorem III.9. *Let $\mathcal{G}$ and $\mathcal{H}$ are two IFGs, then*

$$\widetilde{DI}(\mathcal{G}\times\mathcal{H}) \geq \max\Big\{\widetilde{DI}(\mathcal{G}), \widetilde{DI}(\mathcal{H})\Big\}$$

Proof: Let S be a $\widetilde{DI}-$set of $\mathcal{G}\times\mathcal{H}$ and X be the $m\Big((\mathcal{G}\times\mathcal{H})-S\Big)$. Consider $S_1 = \{x \mid (x,y)\in S\}$ and $S_2 = \{y \mid (x,y)\in S\}$. Clearly, $S_1 \subseteq V(\mathcal{G})$ and $S_2 \subseteq V(\mathcal{H})$. Similarly, let X_1 and X_2 are two the projections of X onto $\mathcal{G}$ and $\mathcal{H}$, respectively. Let $X_1^* = X_1 - S_1$ and $X_2^* = X_2 - S_2$.
S_1 and X_1^* are two subsets of $V(\mathcal{G})$. As S is a DS, of $\mathcal{G}\times\mathcal{H}$, S_1 must be a DS of $\mathcal{G}$. Also after deleting S_1 from $V(\mathcal{G})$, X_1^* is the component with maximum order. Hence $\widetilde{DI}(\mathcal{G}) \leq |S_1| + |X_1^*|$. Same arguments hold for the IFG $\mathcal{H}$. Hence $\widetilde{DI}(\mathcal{H}) \leq |S_2| + |X_2^*|$. Thus,

$$\begin{aligned}\widetilde{DI}(\mathcal{G}\times\mathcal{H}) = |S| &\geq \max\{|S_1|+|X_1^*|, |S_2|+|X_2^*|\}\\ &\geq \max\{\widetilde{DI}(\mathcal{G}), \widetilde{DI}(\mathcal{H})\}.\end{aligned}$$

■

Theorem III.10. *If $\mathcal{G}$ and $\mathcal{H}$ are two IFGs, then*

$$\widetilde{DI}(\mathcal{G}\times\mathcal{H}) \leq |S_1||V(\mathcal{H})| + |S_2||V(\mathcal{G})| - |S_1||S_2| + |X_1||X_2|.$$

Proof: Consider $V(\mathcal{G}) = S_1 \cup X_1 \cup Y_1$ and $V(\mathcal{H}) = S_2 \cup X_2 \cup Y_2$, where $\widetilde{DI}(\mathcal{G}) = |S_1| + |X_1|$ and $\widetilde{DI}(\mathcal{H}) = |S_2| + |X_2|$. Now consider the set $S = \Big(S_1 \times V(\mathcal{H})\Big) \cup \Big(S_2 \times V(\mathcal{G})\Big)$. Since S_1 and S_2 are $\widetilde{DI}-$sets of $\mathcal{G}$ and $\mathcal{H}$, then $S_1 \times S_2$ is a DS. It is easy to verify that S is a DS of $\mathcal{G}\times\mathcal{H}$. Consider $(\mathcal{G}\times\mathcal{H}) - S$. Since X_1 and X_2 are the maximum components of $\mathcal{G}$ and $\mathcal{H}$ after the deletion of S_1 and S_2, respectively. We have $X_1 \times X_2$ as the maximum component of $(\mathcal{G}\times\mathcal{H}) - S$. Hence,

$$\begin{aligned}\widetilde{DI}(\mathcal{G}\times\mathcal{H}) &\leq |S| + |X_1 \times X_2|\\ &= |S_1||V(\mathcal{H})| + |S_2||V(\mathcal{G})| - |S_1||S_2| + |X_1||X_2|.\end{aligned}$$

■

Definition III.11. *The domination weak integrity of an IFG $\mathcal{G} = (A,B)$ is defined as $\min\{|S| + m_e(\mathcal{G}-S) \mid S \subseteq V\}$, where S is an DS of $\mathcal{G}$ and $m_e(\mathcal{G}-S)$ is the order of the largest component in $\mathcal{G}-S$, respectively, and denoted by $\widetilde{DI_w}(\mathcal{G})$. A subset S is a $\widetilde{DI_w}(\mathcal{G})-$set if $\widetilde{DI_w}(\mathcal{G}) = \min\{|S| + m_e(\mathcal{G}-S)\}$ where S a DS of $\mathcal{G}$.*

Remark III.12. *If $\mathcal{G} = (A,B)$ is a connected IFG, then*
i) $\widetilde{DI}(\mathcal{G}) \leq \widetilde{DI_w}(\mathcal{G})$.
ii) $\widetilde{DI_w}(\mathcal{H}) \leq \widetilde{DI_w}(\mathcal{G})$, *if $\mathcal{H}$ is a subgraph of $\mathcal{G}$.*

Example III.13. *Consider an IFG $\mathcal{G} = (A,B)$ as shown in "Fig. 2"*
All edges are strong. The minimal DSs are as follow

$$S_1 = \{a,c,d,f\}, \quad S_2 = \{a,e\}, \quad S_3 = \{b,f\}$$
$$|S_1| = 2.5, \quad |S_2| = 1.6, \quad |S_3| = 1.3$$
$$m_e(\mathcal{G}-S_1) = 0.75, m_e(\mathcal{G}-S_2) = 1.2, m_e(\mathcal{G}-S_3) = 1.2$$

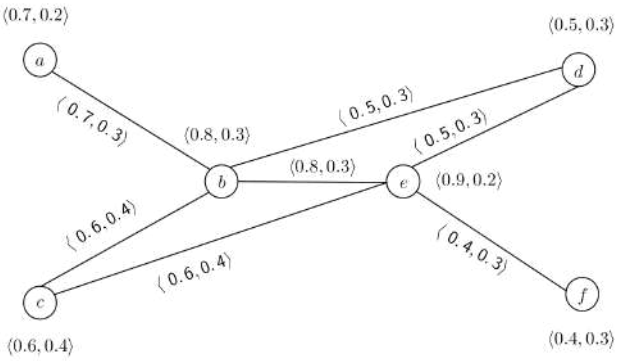


Fig. 2. The Strong IFG $\mathcal{G} = (A,B)$

$$\begin{aligned}\widetilde{DI_w}(\mathcal{G}) &= \min\{2.5+0.75, 1.6+1.2, 1.3+1.2\}\\ &= \min\{3.25, 2.8, 2.5\}\\ &= 2.5\end{aligned}$$

Therefor, S_3 is a $\widetilde{DI_w}-$set.

Theorem III.14. *For any IFG $\mathcal{G} = (A,B)$ and $x \in V$,*

$$\widetilde{DI_w}(\mathcal{G}-x) \geq \widetilde{DI_w}(\mathcal{G}) - |\{x\}|$$

Proof: Let S be a $\widetilde{DI_w}-$set of an IFG $\mathcal{G}-x$. Then, S is a DS of $\mathcal{G}-x$ and $\widetilde{DI_w}(\mathcal{G}-x) = |S| + m_e\Big((\mathcal{G}-x)-S\Big)$. Let $T = S \cup \{x\}\cdot T$ be a DS of $\mathcal{G}$. Therefore,

$$\begin{aligned}\widetilde{DI_w}(\mathcal{G}) &= |T| + m_e(\mathcal{G}-T)\\ &= |S\cup\{x\}| + m_e\Big(\mathcal{G} - (S\cup\{x\})\Big)\\ &= |S| + m_e\Big((\mathcal{G}-x)-S\Big) + |\{x\}|\\ &= \widetilde{DI_w}(\mathcal{G}-x) + |\{x\}|\end{aligned}$$

Thus, $\widetilde{DI_w}(\mathcal{G}-x) \geq \widetilde{DI_w}(\mathcal{G}) - |\{x\}|$.

■

Theorem III.15. *For any IFG $\mathcal{G} = (A,B)$ and $xy \in E$,*

$$\widetilde{DI_w}(\mathcal{G}-x) \geq \widetilde{DI_w}(\mathcal{G}) - \min\{|\{x\}|, |\{y\}|\}$$

Proof: Let S be a $\widetilde{DI_w}-$set of $\mathcal{G}-xy$. Then, S is a DS of $\mathcal{G}-xy$ and $\widetilde{DI_w}(\mathcal{G}-xy) = |S| + m\Big((\mathcal{G}-xy)-S\Big)$. Let $T = S\cup\{x\}$ or $T = S\cup\{y\}$. If x or y belongs to S, then $T = S$. If both x and y do not beloing to S, then $|T| = |S| + |\{x\}|$ or $|T| = |S| + |\{y\}|$. T is a DS of $\mathcal{G}$ and $m_e(\mathcal{G}-T) = m_e\Big((\mathcal{G}-x)-S\Big)$. Therefore,

$$\begin{aligned}\widetilde{DI_w}(\mathcal{G}) &= |T| + m_e(\mathcal{G}-T)\\ &= |S| + \min\Big\{|\{x\}| + |\{y\}|\Big\} + m_e\Big((\mathcal{G}-xy)-S\Big)\\ &= \widetilde{DI_w}(\mathcal{G}-xy) + \min\Big\{|\{x\}| + |\{y\}|\Big\}\end{aligned}$$

Thus, $\widetilde{DI_w}(\mathcal{G}-xy) \geq \widetilde{DI_w}(\mathcal{G}) - \min\Big\{|\{x\}| + |\{y\}|\Big\}$.

■

Remark III.16. *Suppose $xy \in E$ and there exists a $\widetilde{DI_w}-$set S of IFG $\mathcal{G} = (A, B)$ containing x or y. then,*

$$\widetilde{DI_w}(\mathcal{G} - xy) \geq \widetilde{DI_w}(\mathcal{G}).$$

Proposition III.17. *Let $\mathcal{G}$ be a spanning sub graph of an IFG $\mathcal{H}$ and $\widetilde{DI_w}(\mathcal{G}) = \widetilde{DI_w}(\mathcal{H})$. Then, $\widetilde{DI_w}-$set of $\mathcal{G}$ is a $\widetilde{DI_w}-$set of $\mathcal{H}$.*

Theorem III.18. *Let S be a DI_w-set of an IFG $\mathcal{G}$. Then,*

$$m_e(\mathcal{G} - S) \leq \widetilde{DI_w}(\mathcal{G} - S)$$

Proof: Let T be a $\widetilde{DI_w}-$set of $\mathcal{G} - S$

$$\begin{aligned} |S| + m_e(\mathcal{G} - S) = \widetilde{DI_w}(\mathcal{G}) &\leq m_e\Big(\mathcal{G} - (S \cup T)\Big) + |S \cup T| \\ &= |S| + |T| + m_e(\mathcal{G} - S - T) \\ &= |S| + \widetilde{DI_w}(\mathcal{G} - S) \end{aligned}$$

Hence, $m_e(\mathcal{G} - S) \leq \widetilde{DI_w}(\mathcal{G} - S)$ ■

Theorem III.19. *For any IFG $\mathcal{G} = (A, B)$, $\widetilde{DI_w}(\mathcal{G}) \geq \delta(\mathcal{G})$.*

Proof: Let S be $\widetilde{DI_w}-$ set of IFG $\mathcal{G}$ such that $\widetilde{DI_w}(\mathcal{G}) = |S| + m_e(\mathcal{G} - S)$. Therefore,

$$\begin{aligned} m_e(\mathcal{G} - S) &\geq \delta(\mathcal{G} - S) \geq \delta(\mathcal{G}) \\ \widetilde{DI_w}(\mathcal{G}) &= |S| + m_e(\mathcal{G} - S) \\ &\geq |S| + \delta(\mathcal{G}) - |S| \\ &= \delta(\mathcal{G}). \end{aligned}$$

IV. Application

Suppose we want to create a satellite station at different points in the country. Suppose that each region has a station. This can be modeled as a graph theory problem: each region is represented by a vertex. If the station is located in a region, it covers the adjacent regions, then joins together with an edge. Now, we are looking for the dominant set. But we can not trust the dominant set that offers the best model. In times of natural disasters, everyone faces network failure. While some stations break down, we have to cover the maximum area. Therefore, the network provider must select a group of stations that cover the maximum area. This concept is developed as the integrity of the graph. Our new parameter also considers both integrity and domination. Therefore, a control integrity set provides full coverage and ensures a better-fixed network, which covers the maximum area. Depending on the need, the coverage in a particular region is not the same. We face an IFG considering network coverage in an area. Therefore, the domination integrity in IFGs is required.

It consists of six regions with a satellite station. In each region, each vertex has a station. The membership values define the maximum capacity of a station. The edge membership values represent the power of the network signal between the two nodes. The problem of failure of some stations can be investigated using intuitive fuzzy graphs. . The IFG corresponding to the satellite stations, is shown in "Fig. 3".

In this IFG, all edges are strong. The DS S, $m(\mathcal{G} - S)$, $\widetilde{DI}$ values are given in "Table. I". In this table, the lowest value is for the set $\{b, d, f\}$. Therefore, this set assue the reliability and domination property of the network. Thus, the station placed in these nodes gives better coverage even in natural disasters, when compared with other nodes. ■

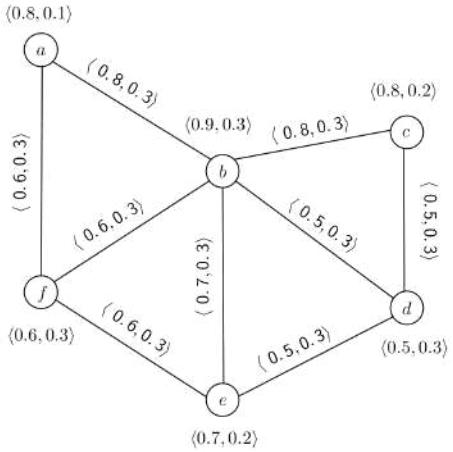


Fig. 3. IFG $\mathcal{G} = (A, B)$ for satellite stations.

TABLE I
Domination integrity of IFG $\mathcal{G} = (A, B)$

S	$m(\mathcal{G} - S)$	$\widetilde{DI}$
$\{b\}$	3.65	4.45
$\{a, d\}$	3	4.45
$\{c, f\}$	3	4.45
$\{d, f\}$	3.2	4.45
$\{b, f\}$	2.15	3.6
$\{b, e\}$	1.5	3.05
$\{b, d\}$	2.25	3.65
$\{a, e, c\}$	2.05	4.45
$\{a, b, d\}$	1.4	3.65
$\{a, b, c\}$	2	4.45
$\{b, d, f\}$	0.85	2.9
$\{a, b, e\}$	1.4	3.8
$\{b, c, e\}$	1.5	3.85
$\{a, c, d, e\}$	1.45	4.45
$\{a, b, c, d\}$	1.4	4.45
$\{a, d, e, f\}$	1.6	4.45
$\{a, c, e, f\}$	1.4	4.45
$\{a, c, d, e, f\}$	0.8	4.45

V. Conclusion

Unlike the connectivity measures, integrity shows not only the difficulty to break down the network but also measures the damage caused. Some of the vulnerability parameters like connectivity, toughness, integrity, binding number etc. are available in the literature. A minimum DS of nodes provides a link with the rest of the nodes in a network, If the removal of such a set, results huge impact in the network. That is, the decision making process is paralyzed but also the communication between the remaining members is minimized. The damage will be more when the DS of nodes are under

attack. In this paper, we studied the concept of domination integrity in IFGs and investigated some of its properties. We also introduced domination weak integrity in the IFG and studied some of the results.

REFERENCES

[1] K. T. Atanassov, "Intuitionistic Fuzzy Sets: Theory and Application," Springer, Physica-Verlag, Heidelberg, Germany, 1999.

[2] S. Ayyasamy, C. Natarajan, "Strong (weak) domination in Fuzzy graphs," International Journal of computational and Mathemtaical secince, 2010.

[3] K. S. Bagga, L. W. Beineke, W. D. Goddard, M. J. Lipman, and R. E. Pippert, "A survey of integrity," Discrete Applied Mathematics, 37/38. 13–28, 1992.

[4] C. A. Barefoot, R. Entringer, and H. Swart, "Vulnerability in graphs-A comparative survey," J. Combin. Math. Combin. Comput, vol. 1, pp. 12–22, 1987.

[5] P. J. Jayalakshmi, C. V. R. Harinarayanan, "Independent and Total strong (weak) domination in Fuzzy Graph", International Journal of computational Engineering Research, 04/Issue, vol. 1, pp. 71–74.

[6] M. G. Karunambigai, R. Parvathi, "Intuitionistic Fuzzy Graphs, Proceedings of 9th Fuzzy Days International Conference on Computational Intelligence," Advances in soft computing: Computational Intelligence, Theory and Applications, Springer-Verlag, vol. 20, pp. 139–150, 2006.

[7] S. Mahde, V. Mathad, "Global domination integrity of graphs," Math Sci Lett, vol. 6, pp. 263—269, 2017.

[8] A. Nagoor Gani, S. Shajitha Begum, "Degree, Order and Size in Intuitionistic fuzzy Graph," International Journal of Algorithm, computing and Mathematics, vol. (3)3, 2010.

[9] A. Nagoor Gani, M. Basher Ahamed, "Order and size in Fuzzy Graphs," Bulletin of pure and Applied Sceinece, vol. 22, pp. 145–148, 2003.

[10] O. Ore, "Theory of graphs," American Mathematical Society, Providence, 1962.

[11] R. Parvathi, G. Thamizhendhi, "Domination in intuitionistic fuzzy graphs," Notes on Intuitionistic Fuzzy Sets, vol. 16, pp. 39–49, 2010.

[12] R. Parvathi, M. G. Karunambigai and K. Atanassov, "Operations on Intuitionistic Fuzzy Graphs," Proceedings of IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 1396–1401, August 2009.

[13] H. Rashmanlou, R. A. Borzooei, "Product vague and its applications," Journal of Intelligent Fuzzy Systems, vol. 30, pp. 371–382, 2016.

[14] H. Rashmanlou, S. Samanta, M. Pal, and R. A. Borzooei, "A study on bipolar fuzzy graphs," Journal of Intelligent Fuzzy Systems, vol. 28, pp. 571–580, 2015.

[15] H. Rashmanlou, R. A. Borzooei, M. Shoaib, Yahya Talebi, M. Taheri, and F. Modnakhaei, "New way for Finding shortest path problem in a Network," J. of Mult.-Valued Logic Soft Computing, vol. 34, pp. 451–460, 2020.

[16] H. Rashmanlou, S. Samanta, M. Pal, and R. A. Borzooei, "Bipolar fuzzy graphs with Categorical properties," International Journal of Computational Intelligence Systems, vol. 8(5), pp. 808–818, 2015.

[17] H. Rashmanlou, S. Samanta, M. Pal, and R.A. Borzooei, "Product of bipolar fuzzy graphs and their degree," International Journal of General Systems, doi.org/10.1080/03081079.2015.1072521, 2015.

[18] A. Rosenfeld, "Fuzzy Graphs," In: L.A. Zadeh, K.S. Fu, M. Shimura (Eds.) Fuzzy Sets and their Applications to Cognitive and Decision Process, Academic Press, New York, pp. 77–95, 1975.

[19] E. Sampathkumar, "The global domination number of a graph," J Math Phys Sci , vol. 23, pp. 377—385, 1989.

[20] M. Saravanan, R. Sujatha, R. Sundareswaran , "A study of regular fuzzy graphs and integrity of fuzzy graphs," Int J Appl Eng Res, vol. 10(82), pp. 160-–164, 2015.

[21] M. Saravanan, R. Sujatha, and R. Sundareswaran, "Concept of integrity and its result in fuzzy graphs," J Intell Fuzzy System, vol. 34(4), pp. 2429—2439, 2018.

[22] A. Shannon, K.T. Atanassov, "A first step to a theory of the intuitionistic fuzzy graphs," in: D. Lakov (Ed.), Proceeding of the FUBEST, Bulgarian Academy of Sciences, Sofia, Bulgaria, pp. 59-–61, 1994.

[23] Sibabrata Ray, S. Jitender, Deogun, "Computational complexity of weighted integrity," J.Comb. Math. Comb. Comput, vol. 16, pp. 65–73, 1994.

[24] A. Somasundaram, S. Somasundaram, "Domination n Fuzzy Graphs-I," Pattern Recognition Letters, vol. 19, 787–791, 1998.

[25] R. Sundareswaran, V. Swaminathan, "Domination integrity in graphs," Proc Int Conf Math Exp Phys Prague, vol. 3(8), pp. 46—57, 2009.

[26] R. Sundareswaran, V. Swaminathan, "Domination integrity of middle graphs," In: T. Chelvam, S. Somasundaram, R. Kala,(eds) Algebra, graph theory and theirapplications. Narosa Publishing House, New Delhi, pp. 88—92, 2010.

[27] R. Sundareswaran, V. Swaminathan, "Domination integrity in trees," Bull Int Math Virtual Inst, vol. 2, pp. 153-–161, 2010.

[28] R. Sundareswaran, V. Swaminathan, "Domination integrity of powers of cycles," Int J Math Res, vol. 3(3), pp. 257–265, 2011.

[29] R. Sundareswaran, V. Swaminathan, "Integrity and domination integrity of gear graphs," J Appl Eng Math, vol. 6(1), pp. 54—64, 2016.

[30] A. A. Talebi, H. Rashmanlou, and S. H. Sadati, "Interval-valued Intuitionistic Fuzzy Competition Graph," J. of Mult-Valued Logic Soft Computing, vol. 34, pp. 335–364, 2020.

[31] A. A. Talebi, H. Rashmanlou, and S. H. Sadati, "New Concepts On m-polar interval-valued intuitionistic fuzzy graph," TWMS J. App. and Eng. Math, vol. 10 (3), pp. 806–818, 2020.

[32] A. A. Talebi, "Cayley fuzzy graphs on the fuzzy group," Computational and Applied Mathematics, vol. 37, 2018.

[33] A. A. Talebi, W. A. Dudek, "Operations on level graphs

of bipolar fuzzy graphs," Bulletin Academiel De Stiinte A Republic Moldova Mathematica, vol. 2 (81), pp. 107–124, 2016.
[34] L. A. Zadeh, "Fuzzy sets," Inf. Control. vol. 8, pp. 338—353, 1965.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)
March 2-4, 2022, Higher Education Complex of Bam

Early COVID-19 Diagnosis from Lung Ultrasound Images Combining RIULBP-TP and 3D-DenseNet

Vida Esmaeili[1], Mahmood Mohassel Feghhi[2], Seyed Omid Shahdi[3]

[1]*Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran,* v.esmaeili@tabrizu.ac.ir
[2]*Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran,* mohasselfeghhi@tabrizu.ac.ir
[3]Faculty of *Electrical, Biomedical and Mechatronics, Qazvin Branch, Islamic Azad University, Qazvin, Iran,* shahdi@qiau.ac.ir

***Abstract*—the pandemic of COVID-19 has affected the world with the high deaths rate. Early diagnosis of this disease is the bottleneck to the patient's health recovery. Its symptoms appear through the wide range of experiments especially accompany with the severe lung lesions. These lesions could be spotted on the lung ultrasound data. Being non-intrusive, low cost, portable, and accurate enough are among the main pros of ultrasound imaging. However, this imaging modality most often contain variety of noises. In order to overcome this challenge, we propose a novel approach combining Rotation Invariant Uniform LBP on 3 Planes (RIULBP-TP) and 3D-DenseNet. These methods are proved to be robust against various noises. Accordingly, our method reaches outstanding results comparing to related most state-of-the-art methods.**

***Index Terms* —COVID-19; lung ultrasound data; RIULBP-TP; 3D-DenseNet**

I. INTRODUCTION

In recent couple of years, the deaths of more than 4,979,421 people have caused widespread panic throughout the world[1]. These deaths and fears are the result of the pandemic coronavirus 2019 (COVID-19). Although more than 6,838,727,352 vaccine doses[1] have been administered in the world, the early diagnosis of this disease still plays a vital role in the patient's health recovery phase [1]. This virus is taking more victims each day due to incomplete vaccination coverage, high level of viral contagiously, and new mutations. Research has shown customary COVID-19 vaccines are roughly 80% effective in preventing the infection [2].

This disease which has been announced to be universal by WHO in 2020, first was spotted inside China in late 2019 [3]. This virus creates wide range of symptoms from mild to serious. High fever, sore throat, dry cough, body ache, and gastrointestinal anomaly have been considered as the main symptoms of COVID-19 [1]. Apparently, in terms of symptoms, it is much like influenza and pneumonia but with higher lung lesions [4]. In severe cases, the symptoms are persistent, where they may be long lasting even beyond for 4 consecutive weeks from the onset [5].

The COVID-19 has caused inadequacy in healthcare resources [6]. During each peak of this pandemic, hospital beds could not come up with rising number of patients. Therefore, immediate and accurate decisions are required for patient admission and discharge. It saves more related resources as well as alleviating the burden on doctors, nurses, and many healthcare workers.

For the diagnostic purpose, reverse transcription-polymerase chain reactions are widely used [7]. However, it has the main drawbacks of relatively poor false-negative and positive rates, time-consuming, and highly reliant on swab location and technique [8]. Furthermore, it is laborious and manual painstaking task.

Another prevalent diagnostic technique is Computed Tomography (CT) scanning images. It is both quick and accurate [9]. Besides, it can provide multiple images from various viewpoints. Nevertheless, this technique is not considered to be fully safe since it radiates x-rays which is very high in energy and hence it is not suitable for infants, pregnant women and vulnerable patients [1]. Moreover, it is not portable and accessible in outskirts and remote cities [1, 9].

X-Ray technique is another choice when it comes to diagnostic task. It generates results much faster than CT [10] as

[1]https://covid19.who.int/?gclid=EAIaIQobChMIsoPM16fw8wIVSqmWCh2swADqEAAYASABEgLsx_D_BwE

it takes just one viewpoint into account. Meanwhile, there is not any further computing post-processing included. Other advantages of X-Ray over CT are its availability and cost effectiveness. However, X-Ray sensitivity is not satisfying when it is compared to the CT [1].

In retrospect, it is critical to apply a diagnostic technique that is simultaneously invasive, precise, portable, cost-effective, and widely available. Interestingly, ultrasound technique satisfies all these requirements [1, 8, and 9]. Nonetheless, it should be mentioned out that ultrasound technique ends up in noisy images which adversely affects the quality [11]. Although advantages of ultrasound techniques outweigh the disadvantage of noisy images, it becomes more fruitful when most of these distractive noises are discarded. This task could be accomplished applying a robust method against noise. Rotation Invariant Uniform Local Binary Pattern (RIULBP or LBP^{riu2}) [12] is one of the most noise insensitive approach to extract textural features [13]. Afterwards, these features should be fed into an optimized learning algorithm to ensure the high accurate early diagnosis.

To this end, we propose combining RIULBP on Three Planes (RIULBP-TP) and 3D-DenseNet method. First, RIULBP applies on a stack of the planes in three orientations on ultrasound image. These planes have been selected due to the direction of B lines and also discontinuous, thick, and unsmooth pleural line. They are specific manifestations of COVID-19 in ultrasound data [14, 15]. The B lines (see Fig. 1) are vertical comet-tail artifacts caused by ultrasound reverberations [1, 14]. It is a marker of disease. These lines are more visible when the air content in the lung is low [15]. Then, the obtained texture maps are fed to the 3D-DenseNet. According to the results, combining RIULBP-TP and 3D-DenseNet provides encouraging results. Fig. 1 depicts the framework of our proposed combining RIULBP-TP and 3D-DenseNet method for early diagnosis of COVID-19 on ultrasound data.

Our main contributions are 1) proposing a noise-free approach applying on ultrasound data 2) improvement of early COVID-19 diagnosis rate compared to the state of arts.

In the following sections related works, proposed method, experimental results, and conclusions are described, in turn.

II. Related Works

A spatial transformer network has been proposed in [8]. Two hundred seventy-seven lung ultrasound data were fed to this network. The average F1 score has been obtained 71.4%. This work was the first experiment of a neural network on COVID-19 ultrasound data. The data have been collected and organized by the authors [8].

Horry et al. [11] have selected an optimized VGG-19 model for applying scarce COVID-19 ultrasound data. They have achieved maximum precision using this model. Their experiment results show that the performance was 0.96 for 2 classes (COVID-19 and Normal) by F1 measurement. Also, they found that the Ultrasound gives the best results compared to X-Ray.

In [16], POCOVID-Net has been suggested for diagnosing COVID-19 on the POCUS [16] which includes 277 bacterial pneumonia, six hundred fifty-four COVID-19, and normal. The POCOVID-Net has included VGG-16, a layer with ReLU activation, an output layer with softmax activation. Its sensitivity, specificity, F1-score, and accuracy are 0.96, 0.79, 0.92, and 89% in 5-fold cross-validation, respectively.

To classify the COVID-19 severity from ultrasound images, MobileNets and also LUSNet [17] based on inceptionV3 have been presented [17-18]. The average accuracy was 95% and 97% using them, correspondingly.

III. Proposed Method

We propose combining RIULBP-TP and 3D-DenseNet for COVID-19 diagnosis. To the best of our knowledge, this is the first work that combines three planes' information with 3D-DenseNet. The planes can seize the noteworthy pixels with main relevant to COVID-19 features. Unlike LBP and its extended versions [19-22], a noise-resistant method is computed on them. Thus, the RIULBP-TP maps are created that are efficient on ultrasound data. These maps are fed to 3D-DenseNet.

Consider a sequence of ultrasound images in a trail. Three planes are supposed on them. The first plane is WH. It can move along the trail. The second plane is LH. A stack of it proceeds on W dimension. The third plane is LW. A stack of LW planes can be thought of along the H dimension. Fig. 2 shows these planes (a) and the stack of them (b).

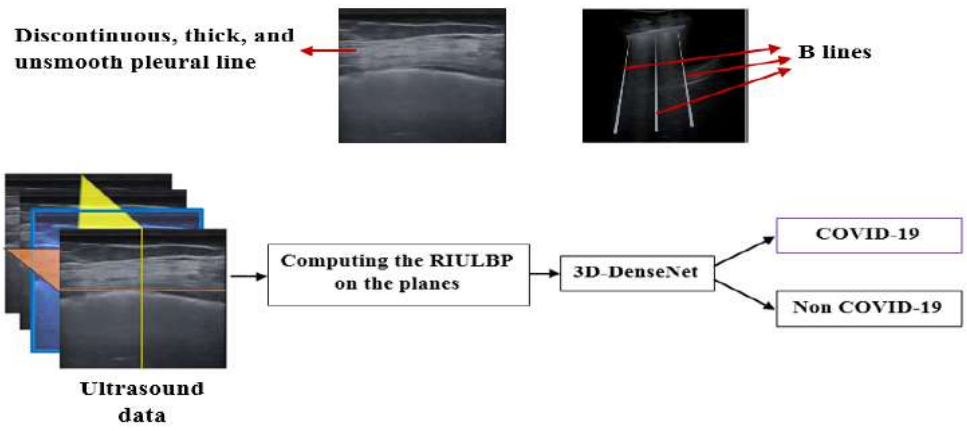


Figure 1. Combining RIULBP-TP and 3D-DenseNet to recognize COVID-19 on ultrasound data.

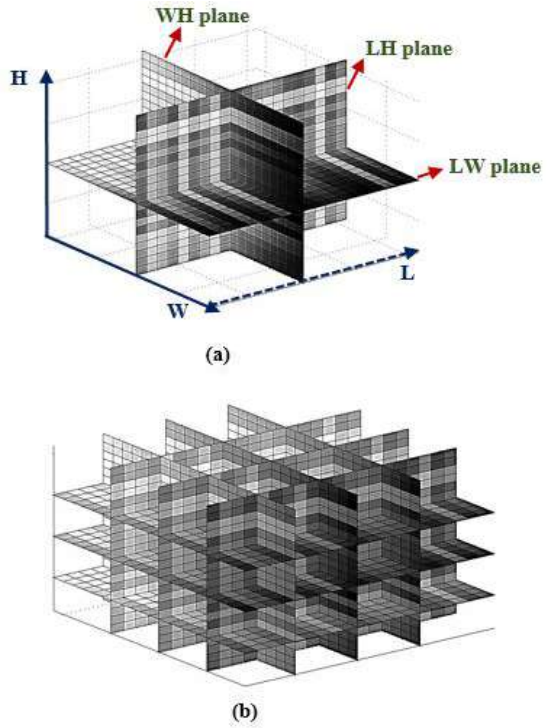


Figure 2. (a) The WH, LW, and LH planes and (b) the stack of them.

Then, the RIULBP is obtained by a circular pixel set (nsp_j) with radius (cra) in the planes. Some of the nsps must be interpolated. The RIULBP can be obtained as follows:

$$RIULBP_{num\ of\ nsps,cra} = \begin{cases} \sum_{j=1}^{num\ of\ nspssig} sig(bvp_j - pccs) & if\ U \le 2 \\ num\ of\ nsps + 1 & otherwise \end{cases} \quad (1)$$

$$sig(bvp_j - pccs) = \begin{cases} 1 & bvp_j \ge pccs \\ 0 & bvp_j < pccs \end{cases} \quad (2)$$

$$U = \left|sig(bvp_{num\ of\ nsps} - pccs) - sig(bvp_1 - pccs)\right| + \sum_{j=1}^{num\ of\ nsps}\left|sig(bvp_j - pccs) - sig(bvp_{j-1} - pccs)\right| \quad (3)$$

Where the *pccs* is the pixel level in the center of a circular symmetric. The *bvp* indicates the brightness value of a pixel. $j=1,\ldots, num\ of\ nsps$.

Afterwards, the created RIULBP maps on the stacks of WH, WL, and HL planes are fed to the 3D-DenseNet [23]. In fact, the temporal information is added to all the internal layers (i.e., 3D pooling, 3D convolution, Rectified Linear Units (ReLU), and fully connected) in DenseNet. The 3D layers learn spatiotemporal features using modeling the temporal information.

For 3D convolution, we add the information along third axis (for instance, temporal dimension) into the kernel, input, and output volume. The input stacks have three (H: height, W: width, and L: length) dimensions. If the input maps are 6*6 by applying zero padding, then a 3*3*3 kernel as a cube is applied on them. Actually, zero padding is to fix the feature size.

In 3D pooling, H*W*L*CM input is taken. CM is the channel of the map. A Hk*Wk*Lk kernel is defined. The Hk, Wk, and Lk are the height, width, and length of the kernel. The stride is defined. Finally, the H_o*W_o*L_o*CM_o output is produced.

Generally, there are 3 dense blocks in the 3D-DenseNet with 40 layers. Before the first dense block, a 3D convolution and kernel is utilized on maps. There is a transition layer between dense blocks. It contains two layers. They are convolution and TL*2*2 pooling. In the first transition, TL is 1. It is 2 on another. The activation function, ReLU, is for increasing the nonlinearity of the decision function and preventing saturating. The softmax has been put in the last layer, e.g., classification. The architecture of the combining RIULBP-TP and 3D-DenseNet is illustrated in Fig.3.

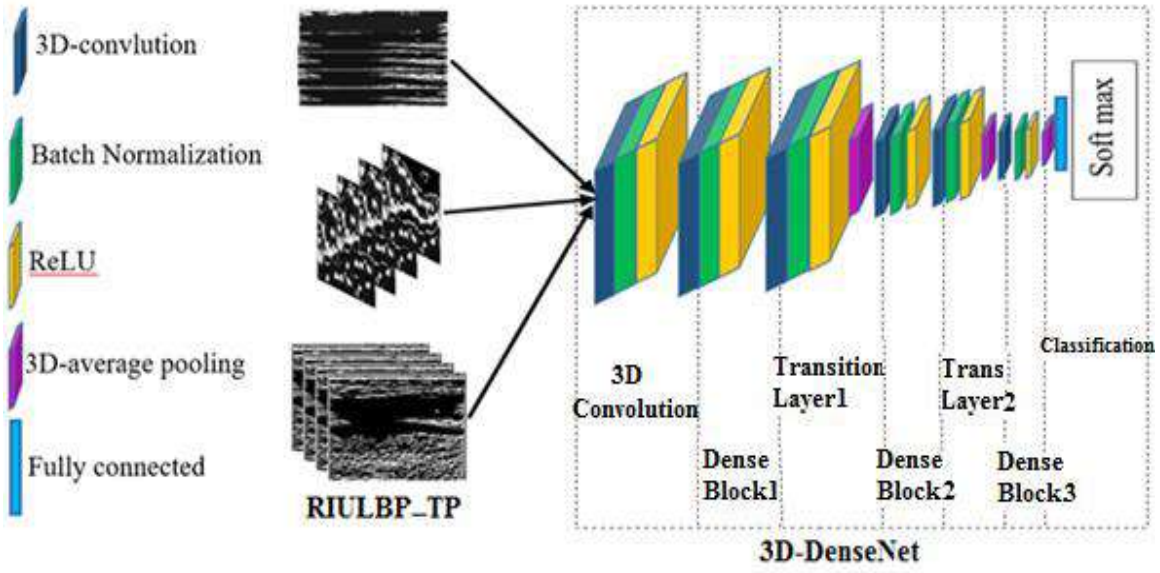


Figure 3. (a) The architecture of the combining RIULBP-TP and 3D-DenseNet

IV. EXPERIMENTAL RESULTS

All experiments are implemented on an Intel i7 3.5GHZ PC. The average training time is 1.5 days. The implementation is done in Python with Tensor-flow.

A. Ultrasound data

We use a public dataset[2] consisting of more than 200 lung ultrasound videos (COVID-19, healthy, and others) [16, 24]. The 115 videos belong to COVID-19, and 75 videos pertain to healthy. This large ultrasound dataset is publicly available[1].

B. Setting parameters

We choose the *number (num) of nsps* = 8, and *cra* = 1. The cause of selecting them is due to the obtained accuracy from the COVID-19 diagnosis. Fig.4 shows the difference in accuracy with several amounts of the*num of nsps* and the *cra*.

The size of the input map has been set to 100x100. For optimizing the network, stochastic gradient descent is utilized. The ratio of test and train ultrasound data is 0.25 (1/4). The batch size and growth rate have been selected 10 and 12, respectively. 0.001 is the initial learning rate at 50 epochs.

C. Results and discussion

The 3D-DenseNet enables the analysis of frames from the ultrasound video. It has good performance when its input has meaningful and robust features. The RIULBP is robust to rotation and noise. Thus, the noise of ultrasound data has no effect on the results. Applying the RIULBP on the stacks of the three planes takes the COVID-19 main features. In fact, the planes can indicate specific manifestations of COVID-19 in ultrasound data.

Table I shows accuracy of combining RIULBP-TP and 3D-DenseNet method for recognition COVID-19 on ultrasound data. In addition, the result of the diagnosis using our proposed combining RIULBP-TP and 3D-DenseNet method has been compared to the state of arts. As can be seen, our proposed method produces a high accuracy. Thus, it can utilize to early recognize the illness from ultrasound data in hospitals and others.

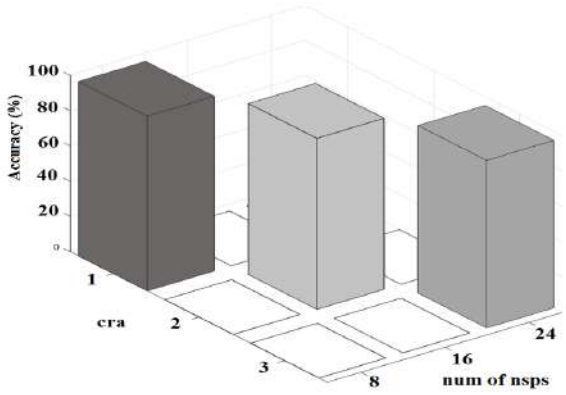


Figure 4. The difference in accuracy with several amounts of the *num of nsps* and the *cra*.

The achieved F1 score (ζ), sensitivity (ψ), specificity (σ), and precision (Γ) are demonstrated in Table II. These metrics are computed as Eq. (4), (5), (6), and (7) [25]. They confirm the promising performance of the proposed method.

$$\psi = \frac{\lambda}{\alpha+\lambda} \tag{4}$$

$$\Gamma = \frac{\lambda}{\beta+\lambda} \tag{5}$$

$$\zeta = 2\left(\frac{\Gamma\times\psi}{\Gamma+\psi}\right) \tag{6}$$

$$\sigma = \frac{\gamma}{\gamma+\beta} \tag{7}$$

where α and β are the false negatives and positives, respectively. Also, the true negatives and positives have been shown by γ and λ.

V. CONCLUSIONS

In this paper, a novel noise-resistant method has been proposed to come up with early diagnosis of COVID19. We take the advantage of selected planes stack to reveal specific symptoms of COVID-19 inside an ultrasound image. According to the comparative results, the proposed combining RIULBP-TP and 3D-DenseNet has provided promising results. Fortunately, it could be further utilized as a diagnostic tool for spotting other similar anomalies within ultrasound data.

TABLE I. COMPARISON OF THE COMBINING RIULBP-TP AND 3D-DENSENET WITH OTHERS FOR COVID-19 RECOGNITION

Method	**Accuracy**
POCOVID-Net [16]	89.1%
MobileNets [18]	95.0%
LUSNet [17]	97.0%
ours	**98.7%**

TABLE II. THE ACHIEVED F1 SCORE (ζ), SENSITIVITY (ψ), SPECIFICITY (σ), AND PRECISION (Γ)

	ζ(%)	ψ(%)	σ(%)	Γ(%)
VGG-19	96	97	98	100
InceptionV3	88	86	89	90
POCOVID-net [16]	92	96	79	87
ours	**99**	**98**	**99**	**100**

[2] https://github.com/jannisborn/covid19_ultrasound

REFERENCES

[1] L. R. Sultan, and C. M. Sehgal, "A review of early experience in lung ultrasound in the diagnosis and management of COVID-19," Ultrasound in Medicine & Biology, vol. 46, no. 9, pp. 2530-2545, 2020.

[2] A. Fowlkes, and et al., "Effectiveness of COVID-19 vaccines in preventing SARS-CoV-2 infection among frontline workers before and during B. 1.617. 2 (Delta) variant predominance—eight US locations, December 2020–August 2021," Morbidity and Mortality Weekly Report, vol. 70, no. 34, p. 1167, 2021.

[3] H. Alabool et al., "Artificial intelligence techniques for containment COVID-19 pandemic: a systematic review," 2020.

[4] S. Rajpal, N. Lakhyani, A. K. Singh, R. Kohli, and N. Kumar, "Using handpicked features in conjunction with ResNet-50 for improved detection of COVID-19 from chest X-ray images," Chaos, Solitons & Fractals, vol. 145, p. 110749, 2021.

[5] A. Nalbandian et al., "Post-acute COVID-19 syndrome," Nature medicine, vol. 27, no. 4, pp. 601-615, 2021.

[6] Y. Zoabi, S. Deri-Rozov, and N. Shomron, "Machine learning-based prediction of COVID-19 diagnosis based on symptoms," npj digital medicine, vol. 4, no. 1, pp. 1-5, 2021.

[7] M. M. Islam, F. Karray, R. Alhajj, and J. Zeng, "A review on deep learning techniques for the diagnosis of novel coronavirus (covid-19)," IEEE Access, vol. 9, pp. 30551-30572, 2021.

[8] S. Roy et al., "Deep learning for classification and localization of COVID-19 markers in point-of-care lung ultrasound," IEEE transactions on medical imaging, vol. 39, no. 8, pp. 2676-2687, 2020.

[9]D. Colombi et al., "Comparison of admission chest computed tomography and lung ultrasound performance for diagnosis of COVID-19 pneumonia in populations with different disease prevalence," European journal of radiology, vol. 133, p. 109344, 2020.

[10] S. Maheshwari, R. R. Sharma, and M. Kumar, "LBP-based information assisted intelligent system for COVID-19 identification," Computers in Biology and Medicine, vol. 134, p. 104453, 2021.

[11] M. J. Horry et al., "COVID-19 detection through transfer learning using multimodal imaging data," IEEE Access, vol. 8, pp. 149808-149824, 2020.

[12] T. Ojala, M. Pietikainen, T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Transactions on pattern analysis and machine intelligence, vol. 24, no. 7, pp. 971-987, 2002.

[13] M. H. Shakoor, and F. Tajeripour. "Repeating average filter for noisy texture classification," Scientia Iranica, vol. 24, no. 3, pp. 1419-1436, 2017.

[14] O. Karakuş et al., "Line artefact quantification in lung ultrasound images of covid-19 patients via non-convex regularization," 2020.

[15] C. McDermott et al, "Sonographic diagnosis of COVID-19: A review of image processing for lung ultrasound," Frontiers in big Data, vol. 4, 2021.

[16] J. Born et al., "POCOVID-Net: automatic detection of COVID-19 from a new lung ultrasound imaging dataset (POCUS)," arXiv preprint arXiv:2004.12084, 2020.

[17] M. R. Panicker, Y. T. Chen, K. V. Narayan, C. Kesavadas, and A. Vinod, "An Approach Towards Physics Informed Lung Ultrasound Image Scoring Neural Network for Diagnostic Assistance in COVID-19," arXiv preprint arXiv:2106.06980, 2021.

[18] A. Almeida et al., "Lung ultrasound for point-of-care COVID-19 pneumonia stratification: computer-aided diagnostics in a smartphone. First experiences classifying semiology from public datasets," in 2020 IEEE International Ultrasonics Symposium (IUS), pp. 1-4: IEEE, 2020.

[19] V. Esmaeili, and S. O. Shahdi, "Automatic micro-expression apex spotting using Cubic-LBP," Multimedia Tools and Applications, vol. 79, no. 27, pp. 20221-20239, 2020.

[20] V. Esmaeili, M. M. Feghhi, and S. O. Shahdi, "Autonomous Apex Detection and Micro-Expression Recognition using Proposed Diagonal Planes," International Journal of Nonlinear Analysis and Applications, vol. 11, pp. 483-497, 2020.

[21] G. Zhao, and M. Pietikainen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," IEEE transactions on pattern analysis and machine intelligence, vol. 29, no. 6, pp. 915-928, 2007.

[22] V. Esmaeili, M. M. Feghhi, and S. O. Shahdi, "Automatic Micro-Expression Apex Frame Spotting using Local Binary Pattern from Six Intersection Planes," arXiv preprint arXiv: 2104.02149, 2021.

[23] D. Gu, "3D densely connected convolutional network for the recognition of human shopping actions," Diss. Université d'Ottawa/University of Ottawa, 2017.

[24] J. Born et al.,"Accelerating detection of lung pathologies with explainable ultrasound image analysis," Applied Sciences, vol. 11, no. 2, p. 672, 2021.

[25] C. Nicholson, "Evaluation metrics for machine learning—accuracy, precision, recall, and F1 defined," ed: Pathmind. http://pathmind.com/wiki/accuracy-precision-recall-f1, 2019.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Efficiency evaluation using fuzzy DEA-TOPSIS and possibility theory

Moosa Darijani [1], Hani Fekri [2]

[1] Technical and Vocational Center, Bam, Iran, moosa.darijani@yahoo.com
[2] Department of Electrical and Computer Engineering, Bam branch, Islamic Azad University, Bam, Iran, fekri@iaubam.ac.ir

***Abstract*—Only possible solution for classical data envelopment analysis (DEA) model is using accurate data. Also, in this model, efficiency evaluation is calculated from the perspective of best relative efficiency. But in evaluations, when we will evaluation, may encounter ambiguous and inaccurate numbers. Also, will measure efficiency from the perspective of the worst relative efficiency. So, in the present study, simultaneously inaccurate data and the best and worst efficiency evaluation approach are used. In the following, the theory of possibility and necessity are proposed to solve the model in optimistic and pessimistic respectively, then, the RC value is used to rank Decision Making Units (DMU's). Finally, a practical example is used to test the model.**

***Index Terms* — evaluation; DEA; TOPSIS; necessity; possibility**

I. INTRODUCTION

A performance evaluation model was first proposed with two inputs and one output by Farrell [1]. Due to, its limitations on inputs and outputs, the model was not successful. Therefore, data envelopment analysis (DEA) model was proposed by Charnes et al[2]. Aims of this model are to measure and compare the relative efficiency of organizational units with multiple inputs and outputs. This reason is a great advantage over the Farrell model. In general, can be said that in this model, decision units (DMU's) are compared with one of the DMU's that has the best possible relative efficiency. If this DMU has best relative efficiency So it introduced as an efficient unit, otherwise it is inefficient.

The classical DEA (CCR) method usually is evaluated in terms of the best possible relative efficiency. If the evaluated DMU is a unit with the best relative efficiency possible, then the DEA is efficient, otherwise it is inefficient. The efficient DMU's always is imagined has perform better than inefficient DMU's in the DEA approach. If the DMU is efficient in this method be evaluated from the term of worst possible relative efficiency and have a lower relative efficiency than the inefficient DMU of this method. Can still say that an efficient DMU performs better than an inefficient DMU? in this situation, the conclusion is quite unclear, Therefore, there is a clear need to combine the best and worst possible relative performance for the overall evaluation of each DMU. Due to the above limitations in the DEA model and the skewed efficiency limitations, to overcome these limitations have proposed a combined DEA-TOPSSIS method for performance evaluation [3]. this model is a good tool for evaluating unit performance. But, when using this model is needed to have accurate input and output information.

But, there are always situations that the inputs and outputs of the units are unclear and ambiguous. Such that when using the DEA is calculated an investment performance evaluation, several parameters need to be estimated without any changes [4, 5]. However, estimating the uncertain parameters in this process is often challenging. Some researchers have proposed various methods to deal with inaccurate and ambiguous data in the DEA[6]. Fuzzy logic and fuzzy sets are one of the methods that can solve the problem of uncertain information. The algebra of fuzzy sets has been discussed to solve These issues by[7]. Researchers have made many efforts to evaluate the performance of uncertain environments and uncertain data, including:

Triantis et al proposed a mathematical programming method by converting fuzzy inputs and outputs to certain numbers and using membership functions[8]. Kao et al Converted Fuzzy Inputs and Outputs to Interval Numbers via $\alpha - cut$and Fuzzy Expansion Principle. Also, The fuzzy data envelopment analysis

model is transformed into a Interval programming model using the CCR approach through $\alpha - cut$ and then in the next step by Saati et al, then they transformed the interval model into a linear programming model through the variable variation process[9]. Soleimani et al designed a fuzzy data envelopment analysis model that can be achieved with the help of specific performance units[10].

In most of the studies reviewed in this paper, can be concluded that in order to access the optimal fuzzy model, these models must to be transformed into a linear or interval DEA model. Therefore, these processes require complex assumptions and calculations, involving; $\alpha - cut$ and variable variation. But in this paper, is referred to the fuzzy DEA-TOPSIS method which includes these studies; in a research is illustrated an ideal FDEA search framework[11] as well as in another study is entitled Welding process selection for repairing nodular cast iron engine block by integrated fuzzy data envelopment analysis and TOPSIS approaches[12]. In these studies, have been used the method of converting uncertain numbers to certain. since, using of fuzzy logic by researchers in DEA models is increasing and this has led to variety use of methods for solving the model, for example the methods that can be applied; theory of possibility and necessity that we use in this study from them. Among the mathematical theories in the field of vague numbers, the possibility theory is considered one of the most appropriate and coherent the methods in the analysis of uncertain decision environments[13]. Following this approach, FDEA models can be transformed into a linear programming model[14, 15].

I. Preliminary and Methodology

A. DEA-TOPSIS Model

Suppose n is the number of DMU's available for evaluation such that each DMU has m inputs and s outputs whose values of inputs and outputs for each DMU_j are denoted by x_{ij} (i = 1، ⋯ ، m) and y_{rj} (r = 1، ⋯ ، s) , respectively. Where all inputs and outputs are known and positive IDMU and ADMU are defined as follows;

Definition 1 – (Ideal Virtual Decision Making Unit) IDMU is a virtual DMU that can be used to generate the lowest number of inputs and maximum outputs.

Definition 2 – (Anti-Ideal Virtual Decision Making Unit) ADMU is a virtual DMU that can be used to generate the maximum inputs and the lowest outputs [3].

Using the above definitions, the input and output values of IDMU and ADMU can be extracted from the inputs and outputs of the DMU's as follows;

$$x_i^{min} = min_j\{x_{ij}\} \quad and \quad x_i^{max} = max_j\{x_{ij}\}. \qquad i = 1. \dots . m$$

$$y_r^{min} = min_j\{yx_{rj}\} \quad and \quad y_r^{max} = max_j\{y_{rj}\}. \quad r = 1. \dots . s \tag{1}$$

Using the above values and the combination of DEA and TOPSIS can be used for a new method of ranking DMU's that It has separation power better than to the CCR method [3] In this model fewer DMU's are efficient.

Calculation of efficiency and ranking each DMU's are done in steps follows:

Step1: the efficiency of IDMU is defined as follows. [3]

$$max \quad \theta_{IDMU} = \frac{\sum_{r=1}^{s} u_r y_r^{max}}{\sum_{i=1}^{m} v_i x_i^{min}} \tag{2}$$

$$s.t \quad \theta_j = \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i \, x_{ij}} \leq 1 \qquad j = 1. \cdots . n$$

$$u_r . v_i \geq \varepsilon \qquad \forall \quad r.i$$

Such that $u_r$and v_i are the decision variables, the values of the output and input , respectively. The value of ε is a positive value that this causes the calculated weights to be chosen the opposite of zero and positive. In the following Using the variables variation Charnes and Cooper, the nonlinear programming of the upper model is transformed into the following linear programming model: [3]

$$max \quad \theta_{IDMU} = \sum_{r=1}^{s} u_r y_r^{max} \tag{3}$$

$$s.t \qquad \sum_{i=1}^{m} v_i x_i^{min} = 1$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i \, x_{ij} \leq 0 \qquad j = 1. \cdots . n$$

$$u_r . v_i \geq \varepsilon \qquad \forall \quad r.i$$

Step2: The best possible relative efficiency of DMU_O is calculated using the fractional programming model for the target decision unit. y_{ro} and x_{io} are outputs and inputs, respectively. While That remains unchanged the value of the best possible relative efficiency of IDMU[3].

$$max \quad \theta_o = \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i \, x_{io}} \tag{4}$$

$$s.to \quad \theta_{IDMU}^{*} = \frac{\sum_{r=1}^{s} u_r y_r^{max}}{\sum_{i=1}^{m} v_i x_i^{min}}$$

$$\frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i \, x_{ij}} \leq 1 \qquad j = 1. \cdots . n$$

$$u_r . v_i \geq \varepsilon \qquad r = 1. \cdots . s \qquad i = 1. \cdots . m$$

The above model can be converted to linear programming model as (3).

Step3: The efficiency of ADMU is defined as follows;[3]

$$min \quad \emptyset_{ADMU} = \frac{\sum_{r=1}^{s} u_r y_r^{min}}{\sum_{i=1}^{m} v_i x_i^{max}} \tag{5}$$

$$s.to \qquad \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i \, x_{ij}} \geq 1 \qquad j = 1. \cdots . n$$

$$u_r . v_i \geq \varepsilon \qquad \forall \quad r.i$$

The above model can be shown as linear programming model as (3).

Step4: The worst possible relative efficiency of DMU_O is computed using the fractional programming model for the target decision unit. y_{ro} and x_{io} are outputs and inputs, respectively. While That remains unchanged the value of the worst possible relative performance of IDMU[3].

$$max \quad \emptyset_{jo} = \frac{\sum_{r=1}^{s} u_r y_{rjo}}{\sum_{i=1}^{m} v_i x_{ijo}} \qquad (6)$$

$$s.to \quad \emptyset^*_{IDMU} = \frac{\sum_{r=1}^{s} u_r y_r^{min}}{\sum_{i=1}^{m} v_i x_i^{max}}$$

$$\frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \geq 1 \qquad j = 1.\cdots.n$$

$$u_r.v_i \geq \varepsilon \qquad r = 1.\cdots.s \qquad i = 1.\cdots.m$$

Model (6) can be converted to linear programming model as (3).

Finally, the value of RC (Relative Close) index is obtained for each decision unit as follows; [3].

$$RC_o = \frac{\emptyset^*_o - \emptyset^*_{ADMU}}{(\emptyset^*_o - \emptyset^*_{ADMU}) - (\theta^*_{IDMU} - \theta^*_O)} \qquad (7)$$

B. Possibility and Necessity DEA model

In recent years, fuzzy set theory has been proposed as a way to quantify ambiguous and inaccurate data in DEA models. Fuzzy DEA (FCCR) models are linear fuzzy programming models that were introduced as fuzzy data CCR models and were presented as follows; [16].

$$(FCCR) \quad min \ \frac{\sum_{r=1}^{s} u_r \tilde{y}_{ro}}{\sum_{i=1}^{m} v_i \tilde{x}_{io}} \qquad (8)$$

$$\frac{\sum_{r=1}^{s} u_r \tilde{y}_{rj}}{\sum_{i=1}^{m} v_i \tilde{x}_{ij}} \leq 1 \qquad j = 1.\cdots.n$$

$$u_r.v_i \geq \varepsilon \qquad \forall \quad r.i$$

Such that u_r and v_i are the decision variables and ε is a positive value similar to CCR model. Also $\tilde{x}_{io}$ are fuzzy inputs that are used by of each DMU_o and $\tilde{x}_{ij}$ are fuzzy inputs for all DMU's. $\tilde{y}_{ro}$ are fuzzy outputs that are produced by of each DMU_o and $\tilde{y}_{rj}$ are fuzzy inputs for all DMU's. In the following, the nonlinear programming of the above model is transformed into linear programming model using variable variation Charnes and Cooper as (3);

Since, primary and dual relation of the CCR model is not obvious for each of fuzzy inputs and outputs. hence the possibility (PCCR) and the necessity (NCCR) approach are provided to calculating the initial FCCR model by [16]. The following Possible Approach (PCCR) is presented;

$$(PCCR) \qquad max \qquad \bar{f} \qquad (9)$$

$$\pi\left(\sum_{r=1}^{s} u_r \tilde{y}_{ro} \leq \bar{f}\right) \geq \beta$$

$$\pi\left(\sum_{i=1}^{m} v_i \tilde{x}_{io} = 1\right) \geq \alpha_\circ$$

$$\pi\left(\sum_{r=1}^{s} u_r \tilde{y}_{rj} - \sum_{i=1}^{m} v_i \tilde{x}_{ij} \leq 1\right) \geq \alpha \qquad j = 1.\cdots.n$$

$$u_r.v_i \geq \varepsilon \qquad \forall \quad r.$$

Also, necessity approach (NCCR) is presented as follows;

$$(NCCR) \qquad max \qquad \bar{f} \qquad (10)$$

$$\mathcal{N}\left(\sum_{r=1}^{s} u_r \tilde{y}_{ro} \leq \bar{f}\right) \geq \beta$$

$$\mathcal{N}\left(\sum_{i=1}^{m} v_i \tilde{x}_{io} = 1\right) \geq \alpha_\circ$$

$$\mathcal{N}\left(\sum_{r=1}^{s} u_r \tilde{y}_{rj} - \sum_{i=1}^{m} v_i \tilde{x}_{ij} \leq 1\right) \geq \alpha \qquad j = 1.\cdots.n$$

$$u_r.v_i \geq \varepsilon \qquad \forall \quad r.i$$

Such that, β and $\alpha_\circ \in [0; 1]$ The possible and necessity levels are defined for constraints (1) and (2) ,respectively. When $\alpha = [\alpha_1.\ldots.\ldots.\alpha_n]^T \in [0.1]^n$ are defined a column vector for the possibility and necessity vector the constraint (3) in the (9)and (10) [16, 17]. In their studies about the theory of possibility and necessity, fuzzy constraints have transformed to certain constraints on the optimal level α for $\tilde{r}_i = (((\tilde{r}_i)^L.(\tilde{r}_i)^M.(\tilde{r}_i)^R))$ as follows;

$$(PCCR) \qquad max \qquad \bar{f} \qquad (11)$$

$$(1-\beta)\left(\sum_{r=1}^{s} u_r \tilde{y}_{ro}\right)^R + \beta\left(\sum_{r=1}^{s} u_r \tilde{y}_{ro}\right)^M \geq \bar{f}$$

$$(1-\alpha_\circ)\left(\sum_{i=1}^{m} v_i \tilde{x}_{io}\right)^L + \alpha_\circ\left(\sum_{i=1}^{m} v_i \tilde{x}_{io}\right)^M \leq 1$$

$$(1-\alpha)\left(\sum_{r=1}^{s} u_r \tilde{y}_{rj}\right)^L + \alpha\left(\sum_{r=1}^{s} u_r \tilde{y}_{rj}\right)^M - (1 - \alpha)\left(\sum_{i=1}^{m} v_i \tilde{x}_{ij}\right)^R - \alpha\left(\sum_{i=1}^{m} v_i \tilde{x}_{ij}\right)^M \leq 1$$

as well as is defined for the theory of necessity as follows; [16, 17]

$$(NCCR) \qquad max \qquad \bar{f} \qquad (12)$$

$$\beta\left(\sum_{r=1}^{s} u_r \tilde{y}_{ro}\right)^L + (1-\beta)\left(\sum_{r=1}^{s} u_r \tilde{y}_{ro}\right)^M \geq \bar{f}$$

$$(1-\alpha_\circ)\left(\sum_{i=1}^{m} v_i \tilde{x}_{io}\right)^M + \alpha_\circ\left(\sum_{i=1}^{m} v_i \tilde{x}_{io}\right)^R \leq 1$$

$$(1-\alpha)\left(\sum_{r=1}^{s} u_r \tilde{y}_{rj}\right)^{M} + \alpha\left(\sum_{r=1}^{s} u_r \tilde{y}_{rj}\right)^{R} - \alpha\left(\sum_{i=1}^{m} v_i\, \tilde{x}_{ij}\right)^{L} - (1-\alpha)\left(\sum_{i=1}^{m} v_i\, \tilde{x}_{ij}\right)^{M} \le 1$$

C. FDEA and *TOPSSIS model*

Suppose, n is the number of decision-making units(DMU's) that each of DMU's have m inputs and s outputs, too the values inputs and outputs are of triangular fuzzy as follows, respectively. $\tilde{x}_{ij} = (x_i^L . x_i^M . x_i^R)$ ، $\tilde{y}_{rj} = (y_r^L . y_r^M . y_r^R)$

Definition 1- according above definition, the values $\tilde{x}_i^{min}$،$\tilde{x}_i^{max}$ ،$\tilde{y}_r^{min}$ ، $\tilde{y}_r^{max}$ are defined as follows. [11]

$$\tilde{x}_i^{min} = \left(\tilde{x}_i^{Lmin} . x_i^{Mmin} . x_i^{Rmin}\right) = \left(\ min_j\{x_{ij}^L\} . min_j\{x_{ij}^M\} . min_j\{x_{ij}^R\}\right) \qquad (13)$$

$$\tilde{y}_r^{min} = (y_r^{Lmin} . y_r^{Mmin} . y_r^{Rmin}) = \left(\ min_j\{y_{rj}^L\} . min_j\{y_{rj}^M\} . min_j\{y_{rj}^R\}\right)$$

$$\tilde{x}_i^{max} = (x_i^{Lmax} . x_i^{Mmax} . x_i^{Rmax}) = \left(\ max_j\{x_{ij}^L\} . max_j\{x_{ij}^M\} . max_j\{x_{ij}^R\}\right)$$

$$\tilde{y}_r^{max} = (y_r^{Lmax} . y_r^{Mmax} . y_r^{Rmax}) = \left(\ max_j\{y_{rj}^L\} . max_j\{y_{rj}^M\} . max_j\{y_{rj}^R\}\right)$$

FCCR calculation are done Similar to CCR model in the following steps:

In the first step; the IDMU efficiency value is calculated for an ideal virtual DMU with fuzzy inputs and outputs and is shown as follows; [11]

$$max \quad \bar{\theta}_{IDMU} = \frac{\sum_{r=1}^{s} u_r \tilde{y}_r^{max}}{\sum_{i=1}^{m} v_i \tilde{x}_i^{min}} \qquad (14)$$

$$s.t \quad \frac{\sum_{r=1}^{s} u_r \tilde{y}_{rj}}{\sum_{i=1}^{m} v_i\, \tilde{x}_{ij}} \le 1 \qquad j = 1. \cdots . n$$

$$u_r . v_i \ge \varepsilon \quad \forall \quad r.i$$

In the second step; The best relative efficiency $\bar{\theta}_o$ is determined for every DMU_O so that IDMU is at the best position and does not change and is defined for each of DMU_O as follows;[9-11]

$$max \quad \bar{\theta}_o = \frac{\sum_{r=1}^{s} u_r \tilde{y}_{ro}}{\sum_{i=1}^{m} v_i\, \tilde{x}_{io}} \qquad (15)$$

$$s.to \quad \frac{\sum_{r=1}^{s} u_r \tilde{y}_r^{max}}{\sum_{i=1}^{m} v_i \tilde{x}_i^{min}} \le \theta_{IDMU}^{*}$$

$$\frac{\sum_{r=1}^{s} u_r \tilde{y}_{rj}}{\sum_{i=1}^{m} v_i\, \tilde{x}_{ij}} \le 1 \qquad j = 1. \cdots . n$$

$$u_r . v_i \ge \varepsilon \qquad r = 1. \cdots . s \qquad i = 1. \cdots . m$$

Such that, $u_r$and v_i are the decision variables, the values of the output and input , respectively. The value of ε is a positive value that is used to be calculated the opposite weights. Also $\tilde{x}_{io}$ are fuzzy inputs that are used by of each DMU_o. $\tilde{y}_{ro}$ are fuzzy outputs that are produced by of each DMU_o. then, variable variation Charnes and Cooper is used to convert the nonlinear programming of the above model to the linear programming model as (3):

In the third step; the ADMU efficiency value is calculated for an anti-ideal virtual DMU with fuzzy inputs and outputs is shown as follows; [11]

$$min \ \bar{\emptyset}_{ADMU} = \frac{\sum_{r=1}^{s} u_r \tilde{y}_r^{min}}{\sum_{i=1}^{m} v_i \tilde{x}_i^{max}} \qquad (16)$$

$$s.t \quad \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i\, x_{ij}} \ge 1 \qquad j = 1. \cdots . n$$

$$u_r . v_i \ge \varepsilon \quad \forall \quad r.i$$

Where similar to (14), the nonlinear programming of the upper model is transformed into the linear programming model using the variables variation Charnes and Cooper as follows;

$$min \quad \sum_{r=1}^{s} u_r \tilde{y}_r^{min} \qquad (17)$$

$$s.to \quad \sum_{i=1}^{m} v_i \tilde{x}_i^{max} = 1$$

$$\sum_{r=1}^{s} u_r \tilde{y}_{rj} - \sum_{i=1}^{m} v_i\, \tilde{x}_{ij} \ge 0 \qquad j = 1. \cdots . n$$

$$u_r . v_i \ge \varepsilon \quad \forall \quad r.i$$

The worst relative efficiency $\bar{\emptyset}_o$ is determined for every DMU_O so that IDMU is at the worst position and does not change and is defined for each of DMU_O as follows; [11].

$$min \ \bar{\emptyset}_o = \frac{\sum_{r=1}^{s} u_r \tilde{y}_{ro}}{\sum_{i=1}^{m} v_i\, \tilde{x}_{io}} \qquad (18)$$

$$s.to \ \frac{\sum_{r=1}^{s} u_r \tilde{y}_r^{min}}{\sum_{i=1}^{m} v_i \tilde{x}_i^{max}} \ge \emptyset_{ADMU}^{*}$$

$$\frac{\sum_{r=1}^{s} u_r \tilde{y}_{rj}}{\sum_{i=1}^{m} v_i\, \tilde{x}_{ij}} \ge 1 \qquad j = 1. \cdots . n$$

$$u_r . v_i \ge \varepsilon \qquad r = 1. \cdots . s \qquad i = 1. \cdots . m$$

Similarly (15), $\tilde{x}_{io}$and $\tilde{y}_{ro}$ are fuzzy inputs and fuzzy outputs for each DMU_o respectively. the nonlinear programming of the above model is transformed into linear programming model with using variable variation Charnes and Cooper .

Finally, after specifying $\bar{\theta}_{IDMU}$, $\bar{\emptyset}_{ADMU}$, $\bar{\theta}_o$ and $\bar{\emptyset}_o$, the relative closeness index of the ideal distance (RC) is calculated as follows;[11]

$$RC_o = \frac{\bar{\emptyset}_o - \bar{\emptyset}_{ADMU}}{((\bar{\theta}_{IDMU} - \bar{\theta}_o) + (\ \bar{\emptyset}_o - \bar{\emptyset}_{ADMU}))} \qquad (19)$$

II. HYBRID MODEL OF FDEA AND TOPSIS USING THEORY OF POSSIBILITY AND NECESSITY

A. Hybrid model of FDEA and TOPSIS using theory of possibility

IDMU Virtual Unit efficiency

If the fuzzy inputs and outputs are normal and convex, using the (9), the PCCR model of the best efficiency is represented for IDMU whit $\bar{\theta}_{IDMU}$ and fuzzy triangular inputs and outputs and using the theory of possibility relationship (11), the fuzzy PCCR model to the linear model is shown as follows;

$$\max \quad \bar{\theta}_{IDMU} \tag{20}$$

$$(1-\beta)\sum_{r=1}^{s} u_r\,\tilde{y}_r^{Rmax} + \beta\sum_{r=1}^{s} u_r\,\tilde{y}_r^{Mmax} \geq \bar{\theta}_{IDMU}$$

$$(1-\alpha_\circ)\sum_{i=1}^{m} v_i\tilde{x}_i^{Rmin} + \alpha_\circ\sum_{i=1}^{m} v_i\tilde{x}_i^{Mmin} \geq 1$$

$$(1-\alpha)\sum_{r=1}^{s} u_r\,\tilde{y}_{rj}^{L} + \alpha\sum_{r=1}^{s} u_r\,\tilde{y}_{rj}^{M} - (1-\alpha)\sum_{i=1}^{m} v_i\tilde{x}_{ij}^{R} - \alpha\sum_{i=1}^{m} v_i\tilde{x}_{ij}^{M} \leq 0 \quad j = 1.\cdots.n$$

$$u_r.v_i \geq \varepsilon \quad \forall \quad r.i$$

Ideal efficiency DMU_O

if the fuzzy inputs and outputs are normal and convex, using the (12), the PCCR model of the best efficiency represented for DMU_O whit $\bar{\theta}_o$ and fuzzy triangular inputs and outputs and using the theory of possibility relationship (11), the fuzzy PCCR model is shown to the linear model as follows;

$$Max \quad \bar{\theta}_o \tag{21}$$

$$(1-\beta)\sum_{r=1}^{s} u_r\tilde{y}_{ro}^{R} + \beta\sum_{r=1}^{s} u_r\tilde{y}_{ro}^{M} \geq \bar{\theta}_o$$

$$(1-\alpha_0)(\sum_{i=1}^{m} v_i\tilde{x}_{io}^{L}) + (\alpha_0)(\sum_{i=1}^{m} v_i\tilde{x}_{io}^{M}) \leq 1$$

$$(1-\alpha_1)\sum_{r=1}^{s} u_r\tilde{y}_{rj}^{L} + \alpha_1\sum_{r=1}^{s} u_r\tilde{y}_{rj}^{M} - (1-\alpha_1)\sum_{i=1}^{m} v_i\,\tilde{x}_{ij}^{R} - \alpha_1\sum_{i=1}^{m} v_i\,\tilde{x}_{ij}^{M} \leq 0 \quad j = 1.\cdots.n$$

$$(1-\alpha_2)\sum_{r=1}^{s} u_r\tilde{y}_r^{Lmax} + \alpha_2\sum_{r=1}^{s} u_r\tilde{y}_r^{Mmax} - \theta_{IDMU}^{*}\,((1-\alpha_2)\sum_{i=1}^{m} v_i\tilde{x}_i^{Rmin} + \alpha_2\sum_{i=1}^{m} v_i\tilde{x}_i^{Mmin}) \leq 0$$

$$u_r.v_i \geq \varepsilon \quad \forall \quad r.i$$

ADMU Virtual Unit efficiency

If the fuzzy inputs and outputs are normal and convex, using the (9), the PCCR model of the worst efficiency represented for ADMU by $\bar{\emptyset}_{ADMU}$ and for fuzzy triangular inputs and outputs and using the theory of possibility (11), the fuzzy PCCR model is shown as follow linear model;

$$min \quad \bar{\emptyset}_{ADMU} \tag{22}$$

$$(1-\beta)(\sum_{r=1}^{s} u_r\tilde{y}_r^{Lmin}) + \beta(\sum_{r=1}^{s} u_r\tilde{y}_r^{Mmin}) \leq \bar{\emptyset}_{ADMU}$$

$$(1-\alpha_0)(\sum_{i=1}^{m} v_i\tilde{x}_i^{Rmax}) + \alpha_0(\sum_{i=1}^{m} v_i\tilde{x}_i^{Mmax}) \geq 1$$

$$(1-\alpha_1)\sum_{r=1}^{s} u_r\tilde{y}_{rj}^{R} + \alpha_1\sum_{r=1}^{s} u_r\tilde{y}_{rj}^{M} - (1-\alpha_1)\sum_{i=1}^{m} v_i\,\tilde{x}_{ij}^{L} - \alpha_1\sum_{i=1}^{m} v_i\,\tilde{x}_{ij}^{M} \geq 0 \quad j = 1.\cdots.n$$

$$u_r.v_i \geq \varepsilon \quad \forall \quad r.i$$

Anti-Ideal efficiency DMU_O

If the fuzzy inputs and outputs are normal and convex, using the (9), the PCCR model of the worst efficiency represented for DMU_O by $\bar{\emptyset}_o$ and for fuzzy triangular inputs and outputs and using the theory of possibility relationship (11), the fuzzy PCCR model is shown as follow the linear model;

$$max \quad \bar{\emptyset}_o \tag{23}$$

$$(1-\beta)\sum_{r=1}^{s} u_r\tilde{y}_{ro}^{L} + \beta\sum_{r=1}^{s} u_r\tilde{y}_{ro}^{M} \leq \bar{\emptyset}_o$$

$$(1-\alpha_0)(\sum_{i=1}^{m} v_i\tilde{x}_{io}^{R}) + (\alpha_0)(\sum_{i=1}^{m} v_i\tilde{x}_{io}^{M}) \geq 1$$

$$(1-\alpha_1)\sum_{r=1}^{s} u_r\tilde{y}_r^{Rmin} + \alpha_1\sum_{r=1}^{s} u_r\tilde{y}_r^{Mmin} - \emptyset_{ADMU}^{*}((1-\alpha_1)\sum_{i=1}^{m} v_i\tilde{x}_i^{Lmax} + \alpha_2\sum_{i=1}^{m} v_i\tilde{x}_i^{Mmax}) \geq 0$$

$$(1-\alpha_2)\sum_{r=1}^{s} u_r\tilde{y}_{rj}^{R} + \alpha_2\sum_{r=1}^{s} u_r\tilde{y}_{rj}^{M} - (1-\alpha_2)\sum_{i=1}^{m} v_i\,\tilde{x}_{ij}^{L} - \alpha_2\sum_{i=1}^{m} v_i\,\tilde{x}_{ij}^{M} \geq 0 \quad j = 1.\cdots.n$$

$$u_r.v_i \geq \varepsilon \quad \forall \quad r.i$$

B. Combined model of FDEA and TOPSIS using theory of necessity

We use to solve the model TOPSIS-FDEA from theory of necessity similar to theory of possibility that its steps are as follows;

IDMU Virtual Unit efficiency

If the fuzzy inputs and outputs are normal and convex, using the (10), the NCCR model of the best efficiency represented for IDMU by $\bar{\theta}_{IDMU}$ and for fuzzy triangular inputs and outputs and using the theory of necessity relationship (12), the fuzzy NCCR model is shown as follows the linear model;

$$\max \quad \bar{\theta}_{IDMU} \tag{24}$$

$$(1-\beta)\sum_{r=1}^{s} u_r\,\tilde{y}_r^{Mmax} + \beta\sum_{r=1}^{s} u_r\,\tilde{y}_r^{Lmax} \geq \bar{\theta}_{IDMU}$$

$$(1-\alpha_\circ)\sum_{i=1}^{m} v_i\tilde{x}_i^{Mmin} + \alpha_\circ\sum_{i=1}^{m} v_i\tilde{x}_i^{Rmin} \leq 1$$

$$(1-\alpha)\sum_{r=1}^{s} u_r\,\tilde{y}_{rj}^{M} + \alpha\sum_{r=1}^{s} u_r\,\tilde{y}_{rj}^{R} - (1-\alpha)\sum_{i=1}^{m} v_i\tilde{x}_{ij}^{M} - \alpha\sum_{i=1}^{m} v_i\tilde{x}_{ij}^{L} \leq 0 \quad j = 1.\cdots.n$$

$$u_r.v_i \geq \varepsilon \quad \forall \quad r.i$$

Ideal efficiency DMU_O

If the fuzzy inputs and outputs are normal and convex, using the (10), the NCCR model of the best efficiency represented for DMU_O by $\bar{\theta}_o$ and for fuzzy triangular inputs and outputs and using the theory of necessity relationship (12), we can transform the fuzzy NCCR model to the linear model as follows;

$$Max \quad \bar{\theta}_o \tag{25}$$

$$(1-\beta)\sum_{r=1}^{s} u_r\tilde{y}_{ro}^{M} + \beta\sum_{r=1}^{s} u_r\tilde{y}_{ro}^{L} \geq \bar{\theta}_o$$

$$(1-\alpha_0)(\sum_{i=1}^{m} v_i\tilde{x}_{io}^{M}) + (\alpha_0)(\sum_{i=1}^{m} v_i\tilde{x}_{io}^{R}) \leq 1$$

$$(1-\alpha_1)\sum_{r=1}^{s} u_r\tilde{y}_{rj}^{M} + \alpha_1\sum_{r=1}^{s} u_r\tilde{y}_{rj}^{R} - (1-\alpha_1)\sum_{i=1}^{m} v_i\,\tilde{x}_{ij}^{M} - \alpha_1\sum_{i=1}^{m} v_i\,\tilde{x}_{ij}^{L} \leq 0 \quad j = 1.\cdots.n$$

$$(1-\alpha_2)\sum_{r=1}^{s} u_r\tilde{y}_r^{Mmax} + \alpha_2\sum_{r=1}^{s} u_r\tilde{y}_r^{Lmax} - \theta_{IDMU}^{*}\,((1-\alpha_2)\sum_{i=1}^{m} v_i\tilde{x}_i^{Mmin} + \alpha_2\sum_{i=1}^{m} v_i\tilde{x}_i^{Lmin}) \leq 0$$

$$u_r.v_i \geq \varepsilon \quad \forall \quad r.i$$

ADMU Virtual Unit efficiency

If the fuzzy inputs and outputs are normal and convex, using the (10), the NCCR model of the worst efficiency represented for ADMU by $\bar{\emptyset}_{ADMU}$ and for fuzzy triangular inputs and outputs and using the theory of necessity relationship (12) we can transform the fuzzy NCCR model to the linear model as follows;

$$min \quad \bar{\emptyset}_{ADMU} \tag{26}$$

$$(1-\beta)(\sum_{r=1}^{s} u_r\tilde{y}_r^{Mmin}) + \beta(\sum_{r=1}^{s} u_r\tilde{y}_r^{Rmin}) \leq \bar{\emptyset}_{ADMU}$$

$$(1-\alpha_0)(\sum_{i=1}^{m} v_i\tilde{x}_i^{Mmax}) + \alpha_0(\sum_{i=1}^{m} v_i\tilde{x}_i^{Lmax}) \geq 1$$

$$(1-\alpha_1)\sum_{r=1}^{s} u_r\tilde{y}_{rj}^{M} + \alpha_1\sum_{r=1}^{s} u_r\tilde{y}_{rj}^{L} - (1-\alpha_1)\sum_{i=1}^{m} v_i\,\tilde{x}_{ij}^{M} - \alpha_1\sum_{i=1}^{m} v_i\,\tilde{x}_{ij}^{R} \geq 0 \quad j = 1.\cdots.n$$

$$u_r.v_i \geq \varepsilon \quad \forall \quad r.i$$

Anti-Ideal efficiency DMU_O

If the fuzzy inputs and outputs are normal and convex, using the (10), the NCCR model of the worst efficiency represented for DMU_O by $\bar{\emptyset}_o$ and for fuzzy triangular inputs and outputs and using the theory of possibility relationship (12) we can transform the fuzzy NCCR model to the linear model as follows;

$$max \quad \bar{\emptyset}_o \tag{27}$$

$$(1-\beta)\sum_{r=1}^{s} u_r\tilde{y}_{ro}^{M} + \beta\sum_{r=1}^{s} u_r\tilde{y}_{ro}^{R} \leq \bar{\emptyset}_o$$

$$(1-\alpha_0)(\sum_{i=1}^{m} v_i\tilde{x}_{io}^{M}) + (\alpha_0)(\sum_{i=1}^{m} v_i\tilde{x}_{io}^{L}) \geq 1$$

$$(1-\alpha_1)\sum_{r=1}^{s} u_r\tilde{y}_r^{Mmin} + \alpha_1\sum_{r=1}^{s} u_r\tilde{y}_r^{Lmin} - \emptyset_{ADMU}^{*}((1-\alpha_1)\sum_{i=1}^{m} v_i\tilde{x}_i^{Mmax} + \alpha_2\sum_{i=1}^{m} v_i\tilde{x}_i^{Rmax}) \geq 0$$

$$(1-\alpha_2)\sum_{r=1}^{s} u_r\tilde{y}_{rj}^{M} + \alpha_2\sum_{r=1}^{s} u_r\tilde{y}_{rj}^{L} - (1-\alpha_2)\sum_{i=1}^{m} v_i\,\tilde{x}_{ij}^{M} - \alpha_2\sum_{i=1}^{m} v_i\,\tilde{x}_{ij}^{R} \geq 0 \quad j = 1.\cdots.n$$

$$\boldsymbol{u_r.v_i \geq \varepsilon} \quad \forall \quad \boldsymbol{r.i}$$

3-3- After specifying $\overline{\emptyset}_{ADMU}.\bar{\theta}_{IDMU}\ .\bar{\theta}_o\ and\ \overline{\emptyset}_o$ we calculate the relative closeness index (RC_o)using the relation (23) in both the theory of possibility and necessity method.

In this section a practical example is used to test the proposed approach. The data for the study are taken from the article by [12]. These dates were used to Welding process selection for repairing nodular cast iron engine block. These dates are from eleven different welding processes, which each of them has 10 criteria, including two input and eight output criteria. These inputs and outputs are triangular fuzzy numbers. the welding processes were ranked using the proposed fuzzy hybrid DEA-TOPSIS method in the study by [12], In the present example, all fuzzy constraints are assumed that them are in satisfying possibility and necessity levels as $\beta = \alpha_0 = \cdots = \alpha_n$. In the first phase, the example is computed in five steps and using five possible levels (0, 0.25, 0.5, 0.75, 1). Initially, we obtained the IDMU efficiency value for the ideal virtual DMU and shown as $\bar{\theta}_{IDMU}$, In the second step, we determined the best relative efficiency for each of the DMU_O , such that IDMU is in the best position possible and does not change. In the third step, we obtained the ADMU efficiency value for the anti-ideal virtual DMU and is presented as $\overline{\emptyset}_{ADMU}$. In the fourth step, we obtained the worst relative efficiency for each of the DMU_O, So that the ADMU is in worst position possible and does not change,. Finally, after specifying $\overline{\emptyset}_{ADMU}.\bar{\theta}_{IDMU}\ .\bar{\theta}_o\ and\ \overline{\emptyset}_o$ we calculated RC index and presented the ranking results in TableI.

in the second phase, similar to before phase the example is calculated in five steps using the theory of necessity method for five different necessity levels (0,0.25,0,5,0.75,1) and we calculated $\bar{\theta}_{IDMU}\ .\bar{\theta}_o\ .\overline{\emptyset}_o and\ \overline{\emptyset}_{ADMU}$, respectively. Then we calculated RC index for the theory of necessity and presented the ranking results in Table II.

TABLE I. RANKING OF EACH DMU USING OF FDEA-TOPSIS AND THEORY OF POSSIBILITY FOR DIFFERENT α

DMU	Welding process	Possibility Levelα' Ranking				
		$\alpha = 0$	$\alpha = 0.25$	$\alpha = 0.5$	$\alpha = 0.75$	$\alpha = 1$
١	SMAW	8	8	8	6	3
٢	FCAW	5	6	6	8	8
٣	GMAW	4	4	5	5	6
۴	SAW	6	7	7	7	7
۵	GTAW	3	1	1	1	1
۶	PAW	3	1	1	1	1
٧	OFW	1	2	2	2	2
٨	EBW	10	10	10	10	10
٩	LBW	9	9	9	9	9
١٠	Flame spray welding	2	3	3	4	5
١١	Furnace welding	7	5	4	3	4

TABLE II. RANKING OF EACH DMU USING OF FDEA-TOPSIS AND THEORY OF NECESSITY FOR DIFFERENT α

DMU	Welding process	Necessity Levelα' Ranking				
		$\alpha = 0$	$\alpha = 0.25$	$\alpha = 0.5$	$\alpha = 0.75$	$\alpha = 1$
1	SMAW	2	2	2	1	1
2	FCAW	6	6	6	6	6
3	GMAW	3	4	4	4	4
4	SAW	7	7	7	7	7
5	GTAW	1	1	1	2	2
6	PAW	1	1	1	2	2
7	OFW	4	3	3	3	3
8	EBW	8	8	8	8	8
9	LBW	9	9	9	9	9
10	Flame spray welding	5	5	5	5	5
11	Furnace welding	11	11	11	11	11

III. CONCLUSION:

Since in recent years, data envelopment analysis models are used widespread by researchers and in classical models of data envelopment analysis are considered input and output amounts as accurate amounts. But, in the real-world, situations are created that do not exist accurate information about the input and output amounts. Rather, it is not possible to specify precise values for inputs and outputs Therefore, models need to identify the efficiency of each of the DMU's, with account inaccurate data. Nowadays, thousands of articles are published in these cases. These articles are dedicated to solve a sample using existing models or to modify models are solved before. The main reason to introduce of new models in this field is the inefficiency of the previous models. In the present study, we tried to provide a new approach to data envelopment analysis models and presented a suitable model. That it can solve the previous problems in the best manner. Therefore, in this paper, the FDEA-TOPSIS model is used, this has led to that, we extract all possible events in uncertain and inaccurate structures. And simultaneously determine the possible incidence and the possible inconsistencies events in the optimistic and pessimistic states.

References

[1] M. J. Farrell, "The measurement of productive efficiency," Journal of the Royal Statistical Society: Series A (General), vol. 120, no. 3, pp. 253-281, 1957.

[2] A. Charnes, W. W. Cooper, and E. Rhodes, "Measuring the efficiency of decision making units," European journal of operational research, vol. 2, no. 6, pp. 429-444, 1978.

[3] Y.-M. Wang and Y. Luo, "DEA efficiency assessment using ideal and anti-ideal decision making units," Applied mathematics and computation, vol. 173, no. 2, pp. 902-915, 2006.

[4] A. Charnes, W. Cooper, A. Y. Lewin, and L. M. Seiford, "Data envelopment analysis theory, methodology and applications," Journal of the Operational Research society, vol. 48, no. 3, pp. 332-333, 1997.

[5] W. W. Cooper, L. M. Seiford, and K. Tone, Data envelopment analysis: a comprehensive text with models, applications, references and DEA-solver software. Springer, 2007.

[6] W. W. Cooper, K. S. Park, and G. Yu, "IDEA and AR-IDEA: Models for dealing with imprecise data in DEA," Management science, vol. 45, no. 4, pp. 597-607, 1999.

[7] L. A. Zadeh, "Fuzzy sets," Information and Control, vol. 8, no. 3, pp. 338-353, 1965.

[8] K. Triantis and O. Girod, "A mathematical programming approach for measuring technical efficiency in a fuzzy environment," Journal of productivity analysis, vol. 10, no. 1, pp. 85-102, 1998.

[9] S. M. Saati, A. Memariani, and G. R. Jahanshahloo, "Efficiency analysis and ranking of DMUs with fuzzy data," Fuzzy optimization and decision making, vol. 1, no. 3, pp. 255-267, 2002.

[10] M. Soleimani-Damaneh, G. R. Jahanshahloo, and S. Abbasbandy, "Computational and theoretical pitfalls in some current performance

measurement techniques; and a new approach," Applied mathematics and computation, vol. 181, no. 2, pp. 1199-1207, 2006.
[11] A. Hatami-Marbini, S. Saati, and M. Tavana, "An ideal-seeking fuzzy data envelopment analysis framework," Applied Soft Computing, vol. 10, no. 4, pp. 1062-1070, 2010.
[12] S. M. Mirhedayatian, S. E. Vahdat, M. J. Jelodar, and R. F. Saen, "Welding process selection for repairing nodular cast iron engine block by integrated fuzzy data envelopment analysis and TOPSIS approaches," Materials & Design, vol. 43, pp. 272-282, 2013.
[13] D. Dubois and H. Prade, "Possibility theory, probability theory and multiple-valued logics: A clarification," Annals of mathematics and Artificial Intelligence, vol. 32, no. 1, pp. 35-66, 2001.
[14] D. Dubois and H. Prade, "possibility theory: Approach to computerized processing of uncertainty, plennm N. 4," 1988.
[15] D. Dubois and H. Prade, "Possibility theory: qualitative and quantitative aspects," in Quantified representation of uncertainty and imprecision: Springer, 1998, pp. 169-226.
[16] S. Lertworasirikul, S.-C. Fang, J. A. Joines, and H. L. Nuttle, "Fuzzy data envelopment analysis (DEA): a possibility approach," Fuzzy sets and Systems, vol. 139, no. 2, pp. 379-394, 2003.
[17] P. Peykani, E. Mohammadi, A. Emrouznejad, M. S. Pishvaee, and M. Rostamy-Malkhalifeh, "Fuzzy data envelopment analysis: an adjustable approach," Expert Systems with Applications, vol. 136, pp. 439-452, 2019.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)
March 2-4, 2022, Higher Education Complex of Bam

Evaluation multi label fetaure selection for text classification using weighted borda count approach

Mohsen Miri[1], Mohammad Bagher Dowlatshahi[2], Amin Hashemi[3]
Department of Computer Engineering, Faculty of Engineering, Lorestan University, Khorramabad, Iran
[1] miri.mo@fe.lu.ac.ir
[2] dowlatshahi.mb@lu.ac.ir
[3] hashemi.am@fe.lu.ac.ir

***Abstract*—Due to the existence of text data, multi-label (ML) text classification is an essential task in machine learning. Feature selection is an essential and effective preprocess to enhance the learning process. Choosing a Multi-Label Feature Selection (MLFS) algorithm is the most basic, critical, and sensitive choice in ML classification operations. If this choice is based on a criterion, it cannot be attributed to always being sound. Choosing the best algorithm must be evaluated using several different criteria to be examined from different aspects. In this article, we turn the issue into an election and use the Weighted Borda Count method for voting. We do the voting in three stages continuously so that a subset of different features does the voting. In the second stage, voting of different methods is done with six criteria, and each criterion selects the methods in order of priority from the beginning to the end. Voting steps 1 and 2 are performed on eighteen text datasets used. Finally, in the final voting stage, the methods are evaluated and voted on by different text datasets. The final result of the voting in the third stage shows the desired MLFS methods based on their performance from beginning to end. According to the experiments performed and the results obtained, it can be seen that the selection of the algorithm based on several different criteria and considering the overall performance of the algorithm will be better than the selection based on one criterion.**

***Index Terms* — Multi-label feature selection; Text classification; Weighted Borda Count; Voting**

I. INTRODUCTION

With the rise of digital textual data, text classification has become a critical task. Today, as the dimension of the text increases, the number of labels related to the content of the text also increases (each text can be assigned to one or more different contents)[1], [2]. Therefore, for ML content textual data, a ML classifier is required. Text classification generally consists of four parts: feature extraction, feature selection (FS), text display, and text classification [3]. FS is an essential component of text classification since redundant and unrelated features in each dataset either do not affect the output or negatively affect [4], [5]. FS is used to remove these features, reduce computational costs, complexity, learning time, and increase classification accuracy [6]–[9].

Classification performance can be evaluated by three different criteria that are not interchangeable [10]. Since one of the most common criteria for evaluating FS methods is classification performance, many text classification studies have used only one criterion to evaluate their method. Therefore, several different evaluation criteria must be considered to choose a suitable FS method for text classification. Of course, a suitable FS method with significant classification performance may not have good efficiency and stability. Finally, several factors must be considered when choosing a feature selection method.

In this article, for the first time, we evaluated ten MLFS methods for text classification based on several criteria and subsets of different features. In this way, we turned the issue into an election. The election in this article takes place in three consecutive stages, first by a subset of features, then by criteria, and finally by various text datasets, which are discussed in detail in the following sections. In this problem, we use six performance evaluation criteria and ten different text datasets to evaluate different methods.

This article consists of 5 parts, which after the introduction, the requirements and methods and brief explanations about them will be discussed in the second part. Section III presented the proposed method in detail. In sections IV and V, the information needed to perform the experiments and their results are documented and concluded.

II. REQUIREMENTS AND METHODS

This section describes the feature selection methods, single-label, and ML classifier, evaluation criteria, selected evaluation methods, and Weighted Borda Count algorithm.

A. Feature selection

Due to the decreasing dimensions of FS methods in text classification, FS can be defined as an operation that selects at least the relevant text features to optimize the classification operation. Feature selection methods are typically divided into filter, wrapper, and embedded methods. Based on the data used, these methods can also be divided into supervised, semi-supervised, and unsupervised categories [11]. Wrapper methods have more calculations, and filter methods are more accurate than wrapper and embedded methods [12].

Filter-based methods can be divided into global and local categories. In the global method, each feature is scored, regardless of the number of classes. While in the local method, it is given a ML score for each label class. In the global method, each feature takes only one score, then the N top of features are selected as a subset of features. Usually, most feature selection methods are in the global grouping. Filter-based methods can be divided into global and local categories. In the global method, each feature is scored, regardless of the number of classes. While in the local method, it is given a ML score for each label class. In the global method, each feature takes only one score, then the N top of features are selected as a subset of features. Usually, most feature selection methods are in the global grouping [13].

B. Single and multi-label learning

Machine learning is implemented on two categories of single-label and ML data (Figure 1). ML datasets have two or more labels. In this data, each instance has a vector of features as $X_i=(x_{i1},x_{i2},x_{i3}, \ldots,x_{iM})$ and also the labels as $Y_i=(y_{i1},y_{i2},y_{i3}, \ldots,y_{iL})$. M and L show the number of features and the number of labels, respectively[14], [15].

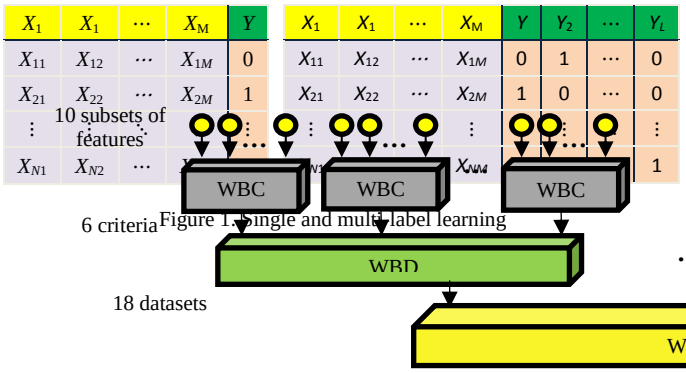


Figure 1. Single and multi-label learning

Suppose $T=\{(x_i,Y_i): i=1,\ldots,n\}$ be a test set, and $L=\{l_j:j=1,\ldots,q\}$ be the set of all labels where $Y_i\subseteq L$, and h be the classifier. According to sample x_i, the labels predicted by the classifier are denoted as Z_i, and $f()$ is a function that is returned by multiple learning systems. We use the following criteria to evaluate the performance of the methods and compare them [14], [17]:

1) *Accuracy:*

The degree of agreement between the measured value and the actual value.

$$Accuracy(h,T)=\frac{1}{n}\sum_{i=1}^{n}\frac{|Y_i\Delta Z_i|}{|L|} \quad (1)$$

2) *Hamming loss:*

The fraction of labels that are incorrectly predicted.

$$Hamming\_loss(h,T)=\frac{1}{n}\sum_{i=1}^{n}\frac{|Y_i\Delta Z_i|}{|L|}, \quad (2)$$

where Δ is the symmetric difference between two sets.

3) *One error:*

one error is the rejection of a true null hypothesis. It measures how many times the ranked label is unrelated to the instance labels.

$$One\_error(f)=\frac{1}{n}\sum_{i=1}^{n}\left[\left[\arg\ \arg\max_{y\in Y_i} f(x_i,y)\right]\right]\notin Y_i \quad (3)$$

4) *Coverage:*

Coverage measures the average number of steps required to cover all relevant labels.

$$Coverage(f)=\frac{1}{n}\sum_{i=1}^{n}\max_{y\in Y_i} rank_f(x_i,y)-1 \quad (4)$$

5) *Ranking loss*

The objective of Ranking Losses is to predict relative distances between inputs.

$$Rank\_loss(f)=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{|Y_i||\bar{Y}_i|}\left|\{(y',y)|f(x_i,y')\leq f(x,y),(y',y'')\in Y_i\times\bar{Y}_i\}\right| \quad (5)$$

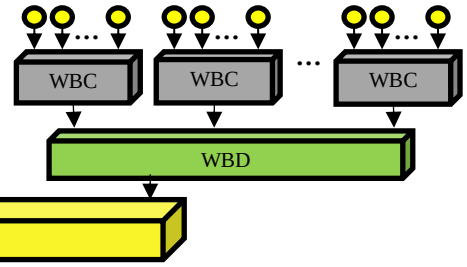


Figure 2. Graphical abstract of the proposed method

C. Evaluation metrics

Performance metrics in classification are fundamental in assessing the quality of learning methods and learned models. However, many different measures have been defined in the literature to make better choices in general or for a specific application area. Evaluation criteria are needed to evaluate the ranking performance of each algorithm. However, researchers have not paid much attention to it [16].

1) *Average precision:*

A measure that combines Recall and Precision measures for ranked retrieval results.

$$Ave\_pre(f)=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{|Y_i|}\sum_{y\in Y_i}\frac{\{y'|\ rank_f(x,y')\leq rank_f(x_i,y),y'\in Y_i\}}{rank_f(x_i,y)} \quad (6)$$

D. Assessed MLFS methods

One of the most critical parts of any classification is feature selection. In this paper, ten different methods for MLFS are used to evaluate and select the appropriate method for text classification. As shown in Table 1, eight of the selected methods are from the filter-based category. Also, MCLS and MDFS methods are part of the embedded group. Filter-based methods can be divided into local and global categories. The eight filter-based methods used in this experiment are placed in the global category.

In selecting algorithms for evaluation, we considered the criteria of the submission time and different structures to have a useful evaluation and comparison. Based on the time criterion, we have selected the superior algorithms that have been introduced since 2019. Also, based on the structure criterion, we have used algorithms with different structures such as population-based [18], Pareto-based, graph-based, reinforcement learning, and based on manifold structure. In addition to these structures, the selected algorithms are placed in multi-objective [9] and single-batch categories. For details on the algorithms, refer to their references in Table 1, and a graphical abstract of the proposed method is presented in Figure 2.

Table 2. Datasets descriptions

Num	Dataset	Samples	Features	Labels
1	Chess	1675	585	227
2	Yelp	10810	671	5
3	Rcv1subset5	6000	1045	101
4	Medical	978	1449	45
5	Philosophy	3971	842	233
6	Chemistry	6961	715	175
7	20NG	19300	1006	20
8	Arts	7484	23150	26
9	Bibtex	7395	1836	159
10	Cooking	10490	577	400
11	Cs	9270	635	274
12	Enron	1702	1001	53
13	Language_log	1460	1004	75
14	Ohsumed	13930	1002	23
15	Recreation	12830	30320	22
16	Science	6428	37190	40
17	Slashdot	3782	1079	22
18	TMC2007	28600	49060	22

Algorithm 1. Weighted Borda Count (WBC)

Input: $K \times F$ Ranks matrix Z
Output: $1 \times K$ Feature ranking vector R
1. $R = \emptyset$;
2. For $i = 1 : K$
3. For $j = 1 : F$
4. $h = Z(i, j)$;
5. $R(j+1) = R(h, 1) + (K - i + 1)$;
6. End
7. End

Table 1. The categorization of multi-label methods

Category	MLFS methods
Filter based	MLACO [19], Ant-TD [20], BMFS [17], MGFS [21], Pareto Cluster [22], LRFS [23], PPT-ReliefF [24], MFS-MCDM [25]
Embedded	MCLS [26], MDFS [27]

E. Weighted Borda Count (WBC)

In short, the WBC is a weight-counting method that gives more importance to the first vote than to the last. Assuming the WBC counts the votes first in each ballot list, each candidate's score is calculated based on the selected priority (Algorithm 1). Then, the total scores are calculated for each candidate, and finally, they are arranged based on their scores.

III. Proposed Method

In this article, we evaluate and select algorithms using the election method. We used eighteen text datasets and six performance evaluation criteria to evaluate ten MLFS methods and select the best method for the text classification. As shown in Fig 2, this election has three consecutive steps in which the voters are subsets of features, criteria, and datasets, respectively. Step by step, the presented method in Algorithm 2 is also visible.

Before the election, ten MLFS methods were implemented on eighteen text datasets that rank the features of each dataset in order of importance (Step 2). In the first step of the election, 10 feature subsets (with 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100 features) vote on MLFS methods based on each evaluation criteria (step 4). The WBC algorithm then counts the available votes. The result of the votes of the subsets of features based on each criterion is allocated to the vote of the same criterion in the second round of elections (step 5). In the second step of elections, six voting criteria send a sequence of MLFS methods as their vote to the WBC algorithm (step 9). The result of this poll is equal to the desired vote for the text dataset on which the MLFS methods have been applied. Finally, after obtaining the final 18 votes of the text datasets, the votes are counted by the WBC, and the MLFS methods are arranged in order of the votes obtained (steps 9 and 10).

IV. Experiments

In this section, experiments and their results are given and discussed. First, the textual datasets used are introduced and explained. Next, the results of the experiments performed are displayed and explained in detail, and discussed.

Algorithm 2. Proposed method

Input: 18 datasets with $M \times N$ feature matrix and $M \times L$ labels matrix
Output: $1 \times M$ MLFS ranking vector F
1. For $i = 1$ to all datasets
2. Run 10 MLFS methods;
3. For $j = 1$ to number of vriteria
4. *SVotes* = 10 feature subsets sort MLFS methods by j-th criterion;
5. *CrVote* (j) = sort (WBC (*SVote*));
6. End
7. *DsVotes* (i) = sort (WBC (*CrVote*));
8. End
9. *FinVotes* = WBC (*DsVotes*)
10. F = sort (FinVotes);

A. Datasets

In the experiments of this empirical study, eighteen ML textual datasets were used. A summary description of the datasets is presented in Table 2. The text datasets used in different dimensions have been selected. For example, the Arts dataset

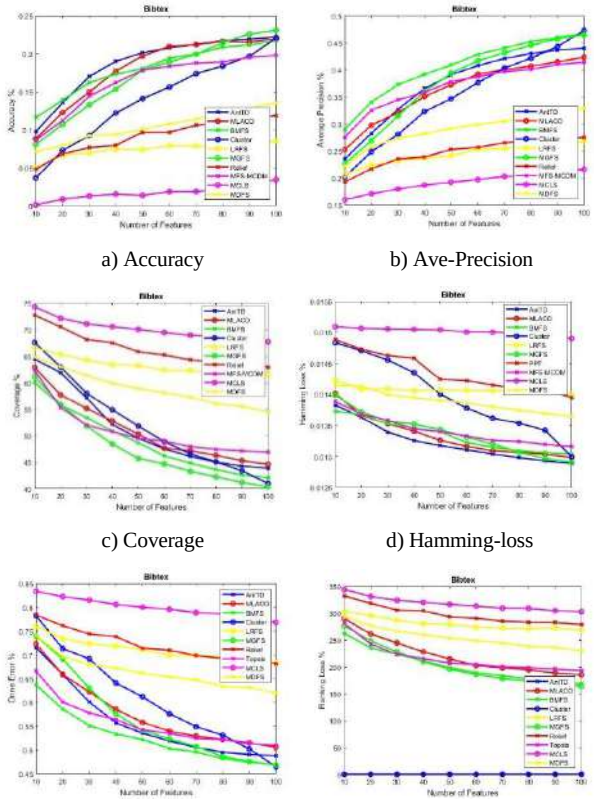
contains 7484 samples with 23150 features, which have very high dimensions.

B. Result

To measure the performance of the MLFS methods, the ML K-nearest neighbor (KNN) classifier [28] with ten neighbors is used, which is compared based on the output of this classifier. At the beginning of the experiments, the data are divided between the training set and the test set with 60% and 40% ratios.

To evaluate, compare, and vote the algorithms used, we used 6 criteria that show the performance of each method on 18 different datasets. Ten equal validations have been used to evaluate this performance of algorithms. Where *m* features are selected as representative from the features extracted from each algorithm, and the test is performed on them. The number of representatives is between 10 and 100 and a factor of 10 ($m \in \{10, 20, 30, 40, 50, 60, 70, 80, 90, 100\}$). All results are calculated based on the average of ten separate runs of the algorithms in different datasets.

After evaluating the methods based on 6 criteria and final voting by datasets, the results are as follows. Figure 3 and Figure 4 show the output information of the MLFS methods in the Bibtex and Arts high-dimensional datasets, which shows the performance of all algorithms based on 6 different criteria.

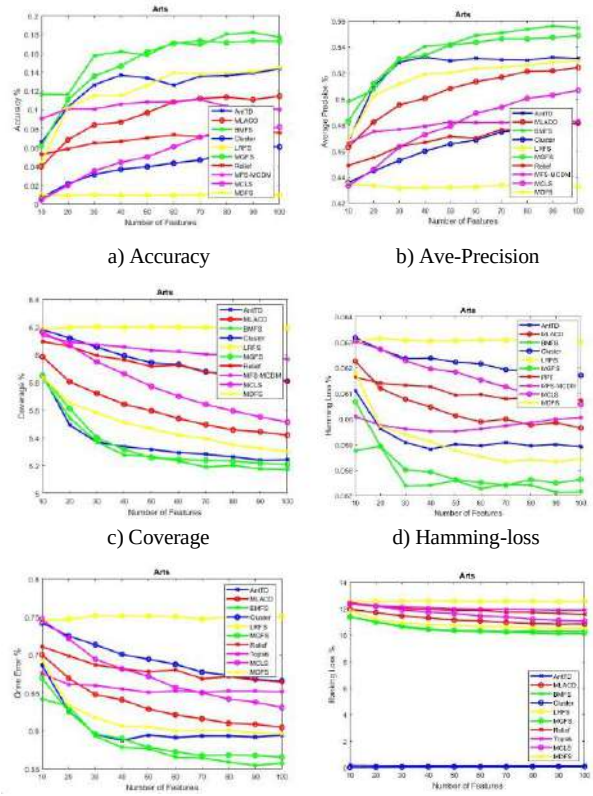


a) Accuracy b) Ave-Precision

c) Coverage d) Hamming-loss

e) One-error d) Ranking-loss

Figure 3. Performance of algorithms on Bibtex dataset



a) Accuracy b) Ave-Precision

c) Coverage d) Hamming-loss

e) One-error d) Ranking-loss

Figure 4. Performance of algorithms on Arts dataset

Table 3. Datasets votes in MLFS methods election

Data Num	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
MALCO	5	5	8	6	5	5	6	4	4	10	4	3	7	4	5	3	8	5
AntTD	1	1	3	3	1	3	2	1	3	1	1	1	1	1	1	1	1	3
BMFS	4	2	5	7	4	7	5	2	7	2	3	5	4	3	3	2	3	6
Pareto Cluster	8	9	7	5	2	6	7	3	8	7	5	4	3	7	8	6	5	8
LRFS	9	10	9	9	8	9	8	8	9	9	6	8	8	10	10	7	7	10
MCLS	10	6	1	1	7	1	1	9	1	6	8	6	10	8	7	10	10	2
MDFS	7	4	2	2	6	2	3	10	2	8	9	7	5	5	4	9	4	1
MGFS	2	3	4	4	3	4	4	6	5	3	2	2	2	2	2	4	2	4
PPT-ReliefF	6	7	10	8	9	8	9	7	6	5	7	9	6	6	6	5	6	7
MFS-MCDM	3	8	6	10	10	10	10	5	10	4	10	10	9	9	9	8	9	9

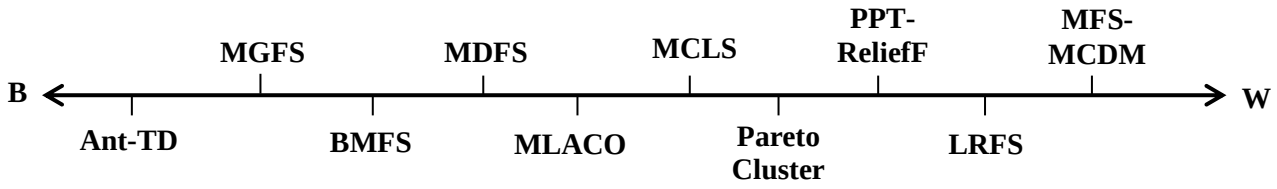


Figure 5. The selected MLFS methods from left to right

Table 3 shows the votes of datasets in the elections, which is related to the third step of the elections. Each row and column of the table corresponds to an MLFS method, and a dataset, respectively, numbered in Table 2. $x(i, j)$ in the table is equal to the priority of selecting the i-th method by the j-th dataset. For example, the Ant-TD method in 12 datasets has been the first choice in voting.

After the election and computation of the final score of the methods based on the votes in Table 3, the MLFS methods used are sorted according to the selection priority, which can be seen in Figure 5 in order from left to right.

V. Conclusion

As the volume of available textual data increases exponentially, the importance of the classification text also increases. As the volume of available ML text data increases exponentially, the importance of the ML text classification also increases. The most critical component of the ML classification is the MLFS. The choice of the MLFS method is an important and sensitive task for classification. Choosing methods based on one criterion can not always lead to a good choice in all respects. This article has evaluated 10 new and powerful MLFS methods based on 6 criteria, 18 text datasets, and 10 subsets of different features. We do this evaluation in the election space in three consecutive steps and use the WBC method to weighted count the votes. First, the feature subsets select the methods based on each criterion. The results of counting the votes of the feature subsets based on each criterion are considered the vote of the same criterion in the next step. Then weighted count of the votes of the criteria is done in each dataset, the result of which is considered the vote of the data set.

Moreover, in the last step, there are 18 votes by datasets, which ultimately chooses the best method. Based on the experiments performed on the used methods, it can be seen that the methods are superior to each other, and the election results will be a proof of these advantages. This article has selected the Ant-TD method as the best method for text classification based on 6 evaluation criteria and 10 subsets of features, and 18 text datasets.

Acknowledgment

This research has not received any special privileges and financial and non-financial assistance from various organiztions.

References

[1] J. Lee and D. W. Kim, "SCLS: Multi-label feature selection based on scalable criterion for large label set," *Pattern Recognit.*, vol. 66, pp. 342–352, Jun. 2017, doi: 10.1016/j.patcog.2017.01.014.

[2] J. Lee, I. Yu, J. Park, and D. W. Kim, "Memetic feature selection for multilabel text categorization using label frequency difference," *Inf. Sci. (Ny).*, vol. 485, pp. 263–280, Jun. 2019, doi: 10.1016/j.ins.2019.02.021.

[3] Kowsari, Jafari Meimandi, Heidarysafa, Mendu, Barnes, and Brown, "Text Classification Algorithms: A Survey," *Information*, vol. 10, no. 4, p. 150, Apr. 2019, doi: 10.3390/info10040150.

[4] M. B. Dowlatshahi, V. Derhami, and H. Nezamabadi-pour, "A novel three-stage filter-wrapper framework for miRNA subset selection in cancer classification," *Informatics*, vol. 5, no. 1, 2018, doi: 10.3390/informatics5010013.

[5] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-Pour, "Gravitational Search Algorithm: Theory, Literature Review, and Applications," *Handb. AI-based Metaheuristics*, pp. 119–150, 2021.

[6] J. Cai, J. Luo, S. Wang, and S. Yang, "Feature selection in machine learning: A new perspective," *Neurocomputing*, vol. 300, pp. 70–79, Jul. 2018, doi: 10.1016/j.neucom.2017.11.077.

[7] A. Hashemi and M. B. Dowlatshahi, "MLCR: a fast multi-label feature selection method based on K-means

and L2-norm," in *2020 25th International Computer Conference, Computer Society of Iran (CSICC)*, 2020, pp. 1–7.

[8] H. Bayati, M. B. Dowlatshahi, and M. Paniri, "MLPSO: a filter multi-label feature selection based on particle swarm optimization," in *2020 25th International Computer Conference, Computer Society of Iran (CSICC)*, 2020, pp. 1–6.

[9] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-pour, "VMFS: A VIKOR-based multi-target feature selection," *Expert Syst. Appl.*, vol. 182, p. 115224, Nov. 2021, doi: 10.1016/j.eswa.2021.115224.

[10] K. N. Junejo, A. Karim, M. T. Hassan, and M. Jeon, "Terms-based discriminative information space for robust text classification," *Inf. Sci. (Ny).*, vol. 372, pp. 518–538, 2016.

[11] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-pour, "A pareto-based ensemble of feature selection algorithms," *Expert Syst. Appl.*, vol. 180, p. 115130, 2021.

[12] J. Li *et al.*, "Feature Selection: A Data Perspective," *ACM Comput. Surv.*, vol. 50, 2017, doi: 10.1145/3136625.

[13] Ş. Taşcı and T. Güngör, "Comparison of text feature selection policies and using an adaptive framework," *Expert Syst. Appl.*, vol. 40, no. 12, pp. 4871–4886, 2013.

[14] S. Kashef, H. Nezamabadi-pour, and B. Nikpour, "Multilabel feature selection: A comprehensive review and guiding experiments," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 8, p. e1240, 2018, doi: 10.1002/widm.1240.

[15] A. Hashemi, M. Bagher Dowlatshahi, and H. Nezamabadi-pour, "An efficient Pareto-based feature selection algorithm for multi-label classification," *Inf. Sci. (Ny).*, vol. 581, pp. 428–447, Dec. 2021, doi: 10.1016/j.ins.2021.09.052.

[16] A. K. Uysal, "An improved global feature selection scheme for text classification," *Expert Syst. Appl.*, vol. 43, pp. 82–92, 2016.

[17] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-Pour, "A bipartite matching-based feature selection for multi-label learning," *Int. J. Mach. Learn. Cybern.*, vol. 12, no. 2, pp. 459–475, Feb. 2021, doi: 10.1007/s13042-020-01180-w.

[18] M. B. Dowlatshahi, V. Derhami, and H. Nezamabadi-Pour, "Fuzzy particle swarm optimization with nearest-better neighborhood for multimodal optimization," *Iran. J. Fuzzy Syst.*, vol. 17, no. 4, pp. 7–24, 2020.

[19] M. Paniri, M. B. Dowlatshahi, and H. Nezamabadi-pour, "MLACO: A multi-label feature selection algorithm based on ant colony optimization," *Knowledge-Based Syst.*, vol. 192, p. 105285, Mar. 2020, doi: 10.1016/j.knosys.2019.105285.

[20] M. Paniri, M. B. Dowlatshahi, and H. Nezamabadi-pour, "Ant-TD: Ant colony optimization plus temporal difference reinforcement learning for multi-label feature selection," *Swarm Evol. Comput.*, vol. 64, p. 100892, Jul. 2021, doi: 10.1016/j.swevo.2021.100892.

[21] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-pour, "MGFS: A multi-label graph-based feature selection algorithm via PageRank centrality," *Expert Syst. Appl.*, vol. 142, p. 113024, 2020.

[22] S. Kashef and H. Nezamabadi-pour, "A label-specific multi-label feature selection algorithm based on the Pareto dominance concept," *Pattern Recognit.*, vol. 88, pp. 654–667, Apr. 2019, doi: 10.1016/j.patcog.2018.12.020.

[23] P. Zhang, G. Liu, and W. Gao, "Distinguishing two types of labels for multi-label feature selection," *Pattern Recognit.*, vol. 95, pp. 72–82, Nov. 2019, doi: 10.1016/j.patcog.2019.06.004.

[24] J. Read, "A pruned problem transformation method for multi-label classification," in *Proc. 2008 New Zealand Computer Science Research Student Conference (NZCSRS 2008)*, 2008, vol. 143150, p. 41.

[25] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-pour, "MFS-MCDM: Multi-label feature selection using multi-criteria decision making," *Knowledge-Based Syst.*, vol. 206, p. 106365, 2020, doi: https://doi.org/10.1016/j.knosys.2020.106365.

[26] R. Huang, W. Jiang, and G. Sun, "Manifold-based constraint Laplacian score for multi-label feature selection," *Pattern Recognit. Lett.*, vol. 112, pp. 346–352, Sep. 2018, doi: 10.1016/j.patrec.2018.08.021.

[27] J. Zhang, Z. Luo, C. Li, C. Zhou, and S. Li, "Manifold regularized discriminative feature selection for multi-label learning," *Pattern Recognit.*, vol. 95, pp. 136–150, Nov. 2019, doi: 10.1016/j.patcog.2019.06.003.

[28] M.-L. Zhang and Z.-H. Zhou, "ML-KNN: A lazy learning approach to multi-label leaming," *Pattern Recognit.*, vol. 40, pp. 2038–2048, 2007, doi: 10.1016/j.patcog.2006.12.019.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Face Recognition based on Multi-shape Morphological Profiles-based Covariance Descriptors and Log- Euclidean Kernel SVM

Behnam Asghari Beirami

Faculty of Geodesy and Geomatics, K. N. Toosi University of Technology, Tehran, Iran, Behnam.asghari1370@gmail.com

***Abstract*—Face recognition (FR) is one of the important areas of image processing that allows us to recognize individuals from gathered face images. It has been shown in the literature that contextual information of the face can improve the performance of FR algorithms. This paper proposes an efficient face recognition method based on multi-shape morphological profiles (MMPs), covariance descriptors, and log-Euclidean kernel SVM. In the first stage of this method, MMPs containing contextual information are generated from the original faces. The covariance descriptors of the MMPs are produced in the subsequent stage. Finally, these covariance descriptors are classified using log-Euclidean kernel SVM. To compare the findings, we ran our tests on ORL face dataset. The mean accuracy of the proposed method is 96.4 on the ORL face datasets, respectively, demonstrating that the proposed method outperforms some of the existing state-of-the-art FR methods.**

***Index Terms* —Face recognition; Covariance descriptor; Morphological profiles; SVM**

I. INTRODUCTION

Face recognition (FR) is a major application of image processing and pattern recognition that has received significant attention in recent years. Face recognition techniques aim to distinguish individuals based on the characteristics of their faces. FR is used in various research disciplines, including human-computer interaction, surveillance, military and criminological applications, and social media management. FR is a challenging problem due to factors such as illumination conditions, head posture, and a limited number of training examples [1].

In general, the two types of feature extraction methods for FR are holistic and local. FR techniques such as eigenface (based on the PCA transform) [2] and Fisherface (based on the LDA transform) [2, 3] , Locally Linear Embedding (LLE) [4, 5], Locally Preserving Projections (LPP) [6, 7], and feature space discriminant analysis (FSDA) [8] are examples of holistic approaches. Due to not modeling local changes in the face, holistic methods are sensitive to illumination condition, expression, and occlusion [9]. In the group of local feature approaches, local binary patterns (LBP) and its family [10, 11], Gabor filters [12, 13], fractal features [14], GLCM, morphological features [8], and deep features [15, 16] can be regarded as the most important methods. A FR method based on the LBP histogram of the face's small regions and the nearest neighbor classifier with chi-square dissimilarity measure is proposed in [17]. According to a study [18], local binary patterns (LBP) and interpolation-based directional wavelet transform (DIWT) are used for FR. For small sample FR with pose variation, Gabor encoding features are combined with nonlinear deep features [19]. A new FR method by combining Gabor wavelets, feature space transformations, and the nearest neighbor classifier is proposed in [12]. A morphological profile (MP) of the face that contains contextual information is combined with the FSDA feature extraction method for FR in [8]. An ensemble deep-learning system based on multiple Gabor face representations has been proposed for FR in [13].

When covariances between local features are taken into account the performance of FR technique based on local features can be enhanced. Tuzel et al. introduced region covariance matrices (RCMs) of several image statistics as a new image region descriptor [20]. In another study, the Gabor-based region covariance matrices (GRCMs) method based on pixel locations and Gabor features is proposed for FR [21]. An improved weighted version of GRCM named Gabor-based weighted region covariance matrix (GWRCM) is also proposed in [22]. Recent advances have evaluated the use of covariance matrices of deep features for facial expression recognition [23].

Morphological profiles (MPs) are powerful techniques for obtaining local contextual features. These features have been less used in the field of FR, despite their superior performance in other fields of image classification. This paper offers a simple and efficient FR approach based on covariance descriptors of multi-shape morphological profiles (MMPs) and Log-Euclidean Kernel SVM. MMPs are generated from multi-shape structuring elements (SEs) in the first stage of the proposed method to extract important contextual information from face images. Next, covariance matrices are calculated to model the relationship between the MMPs. Our final step is to use Log-Euclidean Kernel SVM to classify covariances to identify faces.

The following is the paper's structure. Section 2 gives more details about the methods that are proposed. Our experiments and results are analyzed in section 3. The conclusion of this paper can be found in section 4.

II. METHODOLOGY

The flowchart of the proposed FR method (named CM-MPs-LESVM) can be found in Figure 1. In this method, we used a pre-processing stage at the beginning, which included image smoothing with 3×3 median filtering and resizing all original images. In the second stage, the MMP of each face image is generated. The MMP covariance descriptors are generated at the third stage. After this, the covariance descriptors are fed into Log-Euclidean kernel SVMs for recognizing faces.

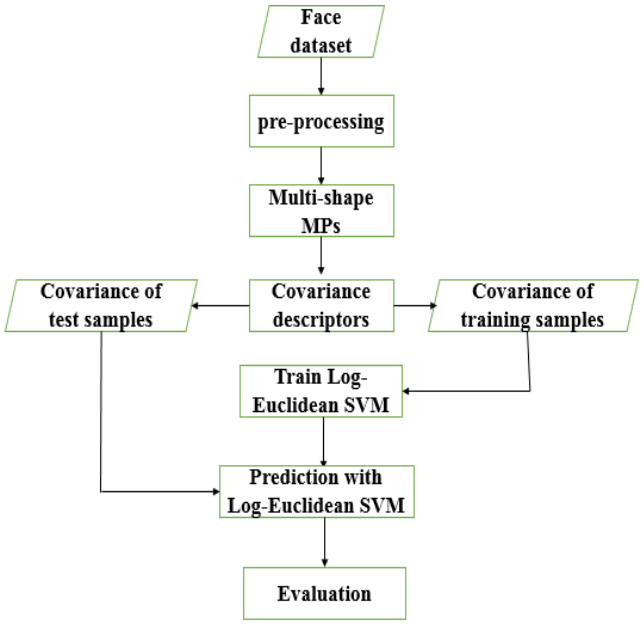


Figure 1. Flowchart of the CM-MMPs-LESVM

A. Multi-shape Morphological profiles

Erosion and dilation are the two basic operators that are used to extract morphological features from greyscale images. Each pixel in the grayscale erosion operator receives the minimum value found in its neighborhood based on SE. In contrast, The highest value determined across the neighborhood of the SE is assigned to each pixel during grayscale dilation. MPs use morphological opening and closing to extract the morphological features. These Morphological operators are defined based on elementary erosion and dilation operators as follows [24]:

$$\text{Opening: } A \circ B = (A \ominus B) \oplus B \quad (1)$$

$$\text{Closing } A \cdot B = (A \oplus B) \ominus B \quad (2)$$

A represents the face image, B represents the SE, $\ominus$ represents the erosion, and $\oplus$ is the dilation. A morphological profile (MP) feature vector of each pixel in coordination with (i, j) defined by the predefined SE consisting of the opening and closing operators are defined as follows [24]:

$$\text{MP}_{(i,j)} = [\, \gamma^{s}_{(i,j)}, \ldots, Q_{(i,j)}, \ldots, \varphi^{s}_{(i,j)} \,] \quad (3)$$

In (3), Q is the original pixel value in coordinate (i,j), φ is closing, γ is opening, and s is the size of the SE. The MP approaches based on fixed shape or fixed-size SEs are inadequate to model all geometric properties of the face. A variety of SE shapes has been shown to be more effective at modeling the geometric properties of faces. In this study, multi-shape morphological profile (MMP) features of the face are generated using SEs with different shapes include disk, diamond, and square. MMP feature vector can be represented by:

$$\text{MMP}_{(i,j)} = [\, \text{MP}^{disk}_{(i,j)}, \text{MP}^{diamond}_{(i,j)}, \text{MP}^{square}_{(i,j)} \,] \quad (4)$$

Where *disk*, *diamond*, and *square* are the shapes of SE. Fig. 2 shows morphological profile features extracted from a sample face image.

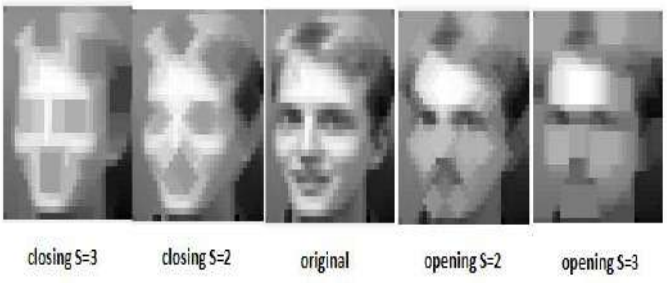


Figure 2. MP of a sample face image with a disk shape SE

B. Covariance descriptor of MMPs

Every extracted morphological feature with a size of $n \times m$ pixels has the same size as the original face image. For each face image, we assumed that the total number of pixels is N ($N = n \times m$) and the total number of MMP features is K. Therefore, final feature vector of each pixel is represented by a $(K+2)$-dimensional feature vector z_i made up of K MMP features and 2 coordinate features (x_i, y_i):

$$z_i = [x_i, y_i, MMP_i^1, MMP_i^2 \ldots, MMP_i^K] \quad (5)$$

As a result, the covariance descriptor (C) of size $(K+2)\times(K+2)$ is computed as follows [21]:

$$C = \frac{1}{N-1}\sum_{i=1}^{N}(z_i - \mu)^T(z_i - \mu) \quad (6)$$

In which μ is the mean vector and T represents the transpose operator. C is a symmetric matrix; whose diagonal elements represent variances and off-diagonal elements represent covariances between features. It is possible to define the

covariance matrix strictly positively by using a regularization technique. In this study, to generate the MMP features, for each SE, the size is considered between 2 and 15. Thus, each SE produces 29 MP features. Assuming three shapes for SE, the final size of z_i in Eq. (5) is 89 (29×3 for MMPs+2 coordinate features).

C. Log-Euclidean Kernel SVM

The produced covariance descriptors (CDs) lie on a manifold by nature and cannot be directly applied to the SVM algorithm. It is possible to resolve this problem by establishing a mapping from Riemannian to Euclidean space using the logm operator. Accordingly, the Log-Euclidean-based kernel function based on the inner product in Euclidean space is defined as [25]:

$$k_{logm}(C1, C2) = trace[logm(C1) \cdot logm(C2)] \qquad (7)$$

It is shown that the proposed Log-Euclidean kernel function Eq (7) meets the conditions of Mercer's theorem and can be used to classify data with SVM [25, 26].

III. Experimental Results

The proposed CM-MMPs-LESVM algorithm was tested on the ORL face database[1]. In the ORL face database, we found 400 images with a resolution of 112 x 92 pixels and 256 gray levels. There are ten images for each of the forty people. In order to capture images of some subjects at different times, a variety of lighting conditions, expressions, and facial details were used. The subjects stood upright, frontally, against a dark, homogeneous background. Fig 3. shows a preview image of the ORL Database.

Figure 3. ORL face database

A cropping operation was performed based on the center of the eyes, and the image was then resized to 32×32 pixels. To make the training set, four, five, and six face images with their labels were chosen at random for each individual. The remaining images regard as the testing set. We run the algorithm fifty times with different sets of samples, and the mean accuracy is reported. In order to measure the effectiveness of the recognition methods, the recognition rate is calculated by dividing the number of correctly classified test samples by the total number of the test samples.

According to Table.1, in the first experiment, the recognition results of covariance descriptors derived from MMPs are compared with those of covariance descriptors derived from other SEs. As a result, disk SE outperforms the other two SEs, and the proposed method of this study based on MMP yielded the best results against the single shape SE.

TABLE I. Comparison of the recognition results of single shape SE MP and MMP

SE	Training size		
	Ts=4	Ts=5	Ts=6
Disk	94.32	96.39	97.09
Diamond	92.09	95.01	95.99
Square	94	95.87	96.47
MMP	**94.83**	**96.77**	**97.60**

In the second experiment, the proposed method (CM-MMPs-LESVM) is compared to the conventional Eigenface method [2], Patch-Based Principal Component Analysis (PPCA) [27], and Multi-resolution dictionary learning (MRDA) [28], improved kernel linear discriminant analysis with probabilistic neural networks (IKLDA-PNN) [29], and supervised kernel locality-based discriminant neighborhood embedding (SKLDNE) [30]. Besides the Eigenface method, which was implemented by the authors, the other results were based on the original papers. To provide a fair comparison, all techniques use the same training and test sample sizes. The means of recognition accuracies of each method are reported in Table 2.

TABLE II. Recognition results of ORL database

Methods	Training size		
	Ts=4	Ts=5	Ts=6
Eigenface	86.19	89.95	91.7
PPCA	---*	91	---
MRDL	---	92.15	---
IKLDA+PNN	91.44	93.95	95.43
SKLDNE	93.33	94	96.87
CM-MMPs-LESVM	**94.83**	**96.77**	**97.60**
***Result is not reported in the original paper**			

According to Table 2, the suggested CM-MMPs-LESVM approach achieves 96.4% mean recognition accuracy across all training sizes, which is approximately 1.66% higher than the nearest competitor method, SKLDNE, which obtains 94.73%. The superiority of CM-MMPs-LESVM is shown by its ability to model the contextual information of the faces with MMPs, along with the use of covariance descriptors to model the relationship between MMPs.

[1] https://cam-orl.co.uk/facedatabase.html

IV. CONCLUSIONS

To recognize faces in this study, we use multi-shape morphological profile-based covariance descriptors and Log-Euclidean Kernel SVM. The study demonstrates that the covariance descriptors of the contextual information produced by MMPs can improve the accuracy of face recognition. On the ORL face databases, experiments have demonstrated that the proposed methods outperform some state-of-the-art methods in the different numbers of the training samples. In the future, we will develop a face recognition method based on the combination of multiple contextual features.

REFERENCES

[1] Singh, S. and S. Prasad, "Techniques and challenges of face recognition: A critical review," Procedia computer science, 143: p. 536-543, 2018.

[2] Turk, M.A. and A.P. Pentland, "Face recognition using eigenfaces. in Proceedings," IEEE computer society conference on computer vision and pattern recognition, 1999.

[3] Anggo, M. and L. Arapu, "Face recognition using fisherface method," in Journal of Physics: Conference Series, 2018. IOP Publishing.

[4] Pang, Y.H., Teoh, A.B.J., Wong, E.K. and Abas, F.S., "Supervised locally linear embedding in face recognition," International Symposium on Biometrics and Security Technologies, 2008. IEEE.

[5] Mekuz, N., C. Bauckhage, and J.K. Tsotsos, "Face recognition with weighted locally linear embedding," in The 2nd Canadian Conference on Computer and Robot Vision (CRV'05), 2005. IEEE.

[6] Wang, R., Nie, F., Hong, R., Chang, X., Yang, X. and Yu, W., "Fast and orthogonal locality preserving projections for dimensionality reduction," IEEE Transactions on Image Processing, Vol. 26(10), pp. 5019-5030, 2017.

[7] Liu, X., L. Shen, and H. Fan, "Face recognition algorithm based on Gabor wavelet and locality preserving projections," Modern Physics Letters B, Vol. 31(19-21), pp. 1740041, 2017.

[8] Imani, M. and G.A. Montazer, "Face recognition using morphological profile and feature space discriminant analysis," in 2017 Iranian Conference on Electrical Engineering (ICEE), 2017. IEEE.

[9] Min, R., A. Hadid, and J.-L. Dugelay, " Improving the recognition of faces occluded by facial accessories," in 2011 IEEE International Conference on Automatic Face & Gesture Recognition (FG), 2011. Santa Barbara, CA, USA: IEEE.

[10] Karanwal, S. and M. Diwakar, "OD-LBP: Orthogonal difference-local binary pattern for Face Recognition," Digital Signal Processing, Vol. 110, pp. 102948, 2021.

[11] Kas, M., El-merabet, Y., Ruichek, Y. and Messoussi, R, "A comprehensive comparative study of handcrafted methods for face recognition LBP-like and non LBP operators," Multimedia Tools and Applications, Vol. 79(1), pp. 375-413, 2020.

[12] Choi, J.Y. and B. Lee, "Ensemble of deep convolutional neural networks with Gabor face representations for face recognition," IEEE Transactions on Image Processing, Vol. 29, pp. 3270-3281, 2019.

[13] Li, L., Ge, H., Tong, Y. and Zhang, Y., "Face recognition using gabor-based feature extraction and feature space transformation fusion method for single image per person problem," Neural Processing Letters, Vol. 47(3), pp. 1197-1217, 2018.

[14] Beirami, B.A. and M. Mokhtarzade, "Spatial-spectral classification of hyperspectral images based on multiple fractal-based features," Geocarto International, pp. 1-15, 2020. https://doi.org/10.1080/10106049.2020.1713232

[15] Wang, M. and W. Deng, "Deep face recognition: A survey," Neurocomputing, Vol. 429, pp. 215-244, 2021.

[16] Ling, H., Wu, J., Huang, J., Chen, J. and Li, P., "Attention-based convolutional neural network for deep face recognition," Multimedia Tools and Applications, Vol. 79(9), pp. 5595-5616, 2020.

[17] J. Aïzan, E.C. Ezin, and a.C. Motamed, "A Face Recognition Approach Based on Nearest Neighbor Interpolation and Local Binary Pattern," in 2016 12th IEEE International Conference on Signal-Image Technology & Internet-Based Systems (SITIS). 2016. Naples, Italy.

[18] Muqeet, M.A. and R.S. Holambe, "Local binary patterns based on directional wavelet transform for expression and pose-invariant face recognition," Applied Computing and Informatics, Vol. 15(2), pp. 163-171, 2019.

[19] Zou, G., Fu, G., Gao, M., Pan, J. and Liu, Z., "A new approach for small sample face recognition with pose variation by fusing Gabor encoding features and deep features," Multimedia Tools and Applications, Vol. 79(31), pp. 23571-23598, 2020.

[20] Tuzel, O., F. Porikli, and P. and Meer, "Region covariance: a fast description for detection and classification," in Proc. European Conf. Computer Vision. 2006. Springer.

[21] Pang, Y., Y. Yuan, and X. Li, "Gabor-based region covariance matrices for face recognition," IEEE Transactions on circuits and systems for video technology, Vol. 18(7), pp. 989-993, 2008.

[22] Qin, H.F., Qin, L., Xue, L. and Yu, C.B., "Gabor-based weighted region covariance matrix for face recognition," Electronics letters, Vol. 48(16), pp. 992-993, 2012.

[23] Otberdout, N., Kacem, A., Daoudi, M., Ballihi, L. and Berretti, S, "Deep covariance descriptors for facial expression recognition," arXiv preprint arXiv:1805.03869, 2018.

[24] Beirami, B.A. and M. Mokhtarzade, "Spatial-Spectral Classification of Hyperspectral Images Based on Extended Morphological Profiles and Guided Filter," Journal of Computer and Knowledge Engineering, Vol. 2(2), 2019. https://doi.org/10.22067/CKE.V2I2.81519

[25] Wang, R., Guo, H., Davis, L.S. and Dai, Q., "Covariance discriminative learning: A natural and efficient approach to image set classification," in 2012 IEEE conference on computer vision and pattern recognition. 2012. IEEE.

[26] Fang, L., He, N., Li, S., Plaza, A.J. and Plaza, J., "A new spatial–spectral feature extraction method for hyperspectral images using local covariance matrix representation," IEEE Transactions on Geoscience and Remote Sensing, Vol. 56(6), pp. 3534-3546, 2018.

[27] Jiang, T.X., Huang, T.Z., Zhao, X.L. and Ma, T.H., "Patch-based principal component analysis for face recognition," Computational intelligence and neuroscience, Vol. 2017, 2017.

[28] Luo, X., Y. Xu, and J. Yang, "Multi-resolution dictionary learning for face recognition," Pattern Recognition, Vol. 93, pp. 283-292, 2019.

[29] Ouyang, A., Liu, Y., Pei, S., Peng, X., He, M. and Wang, Q, "A hybrid improved kernel LDA and PNN algorithm for efficient face recognition," Neurocomputing, Vol. 393, pp. 214-222, 2020.

[30] Khalili Mobarakeh, A., J.A. Cabrera Carrillo, and J.J. Castillo Aguilar, "Robust face recognition based on a new supervised kernel subspace learning method," Sensors, Vol. 19(7), p. 1643, 2019.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Farsi Text Detection and Localization in Videos and Images

Based on YOLO object detection model

Mobina Mosannafat[1], Fatemeh Taherinezhad[2], Hassan Khotanlou[3], Elham Alighardash[4]

[1] *RIV Lab., DEPT. of Computer Engineering, Bu-Ali Sina University, Hamedan, Iran,* mobinamossanafat1999@gmali.com
[2] *RIV Lab., DEPT. of Computer Engineering, Bu-Ali Sina University, Hamedan,* fthgaerin00@gmail.com
[3] *RIV Lab., DEPT. of Computer Engineering, Bu-Ali Sina University, Hamedan,* khotanlou@basu.ac.ir
[4] *RIV Lab., DEPT. of Computer Engineering, Bu-Ali Sina University,Hamedan,* e.alighardash@eng.basu.ac.ir

***Abstract*—Automatic text detection and recognition in images and videos have emerged and aroused widespread interest in recent years due to the dramatic growth of visual information. It seems that there is the lack of any effective model for Farsi text detection in images. In this paper, a new framework is proposed for the Farsi text detection and localization using the up-to-date real-time object detection framework YOLOv5 in videos and images. To evaluate the novel model, a new dataset of news videos is collected. Experimental results show that the proposed model achieves quite promising performance on the new dataset.**

***Index Terms* —text detection; Farsi (Persian) text; YOLO**

I. INTRODUCTION

The fast growth of technology and the use of social media has led to a huge mass of visual data over the internet, in particular, images and videos. Each video frame contains many types of information and among them, texts provide more instinctive information closely related to the content of the video. These text blocks provide information such as location, speaker name, content topic, date and time, and more detailed descriptions. There are four stages to text understanding systems: text detection and text localization, which are essential for system performance, in addition to text extraction and text recognition. Text detection and localization can be used in many daily activities such as reading signs in self-driving cars [1], helping blind persons in navigation [2] [3], and automatic data categorization.

Although text detection in images has become a hot topic in current research, most of the previous works consider the texts in English and Chinese, and yet detection of Farsi text in images has been neglected. Even though Farsi has got four characters more than Arabic, the other 28 characters are the same and they are used in other languages such as Urdu or Pashto. Therefore, systems designed to detect and recognize Farsi texts can also be used for these three languages with a few modifications. Farsi is written from right to left and each word in Farsi may consist of several separated sub words. Farsi alphabets are connected and placed on a hypothetical line called the baseline and there is no upper or lower case in this language. Still, some of the characters can have different shapes according to their position in the word. Also, out of the thirty-two letters, eighteen of them have dots ranging from one to three.

Due to the wide variety of text size, style, orientation, and alignment as well as complex background, lightning conditions, low contrast, and low resolution, accurate text detection in images is yet a difficult task to conduct.

The combination of deep learning with computer vision has proven to be efficient and effective in many Artificial Intelligence (AI) applications. By increasing computing power and the availability of large datasets, it has become possible to train deep neural networks and generate more suitable models. In this paper, we attempt to convert the multi-classification task into binary classification (detecting text from background) and apply it to Farsi text detection and localization in video frames.

The contributions in this study are as follows:

- The text detection part of a novel framework is introduced based on a well-known and efficient detection model.

- A new dataset of videos containing Farsi text is gathered and prepared because of the lack of a proper benchmark for Farsi text detection and recognition.

The rest of this paper is organized as follows. First, we review the related works of the text detection domain. Then, our proposed model and the newly gathered dataset are explained in detail. Next, the experimental results, further analysis, and future directions are presented.

II. RELATED WORKS

A. Traditional text detection methods

Scene text detection and recognition has attracted a lot of attention over the last decade and has become a hot topic in computer vision research. **Traditional** text detection methods first find candidate regions of interest. The handcrafted features are then considered to eliminate non-text regions and to gather text location information. These techniques include methods employing sliding windows (SW) and connected components (CC).

The approach of SW methods for detecting text is a top-down strategy. These methods scan all regions of the image by a multi-scale detection window, select the candidate desired regions, and apply a pre-trained classifier on what is scanned by the sliding window to detect areas that are containing text. The considerable number of redundant sliding windows, which leads to inefficiency, and the inability to detect skewed and curved text lines are some obvious drawbacks of these methods.

CC-based methods, on the other hand, use a top-down strategy and extract components with similar features such as text color, texture, corner map, etc., and then use manually implemented rules or a pre-trained classifier to eliminate non-text areas. In [4], corner map is used to find candidate text regions and a few geometrical rules based on the Farsi language properties are used to filter out non-text areas. In a similar approach, [5] uses corner maps along with texture intensity picture created using discrete cosine transform (DCT) and local binary pattern (LBP) to identify text regions and generate features that result in higher accuracy compared to [4]. Darab et al. [6] used the edge-map constructed by the Sobel operator along with the color properties of the image to extract features. SVM was used to evaluate text regions.

Maximum Stable Extreme Regions (**MSER**) and Stroke Width Transform (**SWT**) are two of the most known methods of CC-based approaches. DRT [7] uses MSER to find candidate text regions and SWT to remove non-textual regions of video frames in the detection phase. Scale-invariant feature transform (SIFT) is then used to track the text in videos while SIFT is used with a bi-cubic interpolation to calculate frequency chart for gradients in [8] to extract features of text areas to detect them. MSER has been reported to be sensitive to image blur and may also produce repetitive components. Chen et al. [9] have used edge-enhanced MSER which is the combination of complementary properties of Canny edges and MSER along with SWT for text detection in low-resolution images. Xu-Cheng et al. [10] proposed two algorithms based on the parent-child relationship of the MSER tree to overcome the repetitive components problem. Although traditional methods have been used widely and achieved acceptable performance in standard databases, they lag behind deep learning approaches in accuracy and speed.

B. Deep learning-based methods

The development of deep learning was a new successful course in text detection. **Deep learning methods** automatically extract features from the image by training a model and are more robust, effective, and simple compared with the traditional methods. Segmentation methods, region proposal-based methods, and hybrid methods are the main categories of deep learning-based text detection methods [11].

Segmentation-based methods segment text from the background at pixel level. Each point in the image is classified as text or non-text and the image is analyzed. Time-consuming post-processing is the principal weakness of these algorithms. Moreover, they can lag detecting sticking or overlapping text. AT-Text [12] proposed a new segmentation-based approach to adopt a model on scene text to provide a semantic mask of text regions for text detection. The deep segmentation model that was applied in this method filter out false-positive components efficiently. Also, to find text regions better and more reliably, MSER is used.

Region proposal-based methods have different classes including two-stage methods and one-stage methods. **Two-stage** methods usually consist of two steps: obtaining region proposals and using classification and regression networks to detect the region proposals. The most representative network is Faster R-CNN [13] which has a similar design to Fast R-CNN [14] . The selective search method is replaced by region proposal network (RPN) in Faster R-CNN which solves the extreme time overhead in generating region of interest. Faster R-CNN is the baseline of many state-of-the-art models like R2CNN [15].

Hybrid methods are a combination of segmentation-based and region-based methods. Segmentation is used to predict the score map and then the regression is used to obtain bounding boxes. These methods support robustness and can manage strings with different forms.

Although these methods are pretty fast, they still cannot meet the real-time requirements. **One-stage** methods can perform the regression and classification task without the need to find region proposals and this will make them faster compared to two-stage methods. Typical one-stage networks are Single Shot Multibox Detector (**SSD**) [16] and You Only Look Once (YOLO) [17]. SSD has a significant advantage of detection speed compared with Faster R-CNN. SSD's architecture builds on the venerable VGG-16 architecture but discards the fully connected layers. The classification and regression tasks are done on the multi-scale feature maps build.

YOLO predicts bounding boxes and multiple objects but compared to SSD, it does not use multi-scale feature maps. YOLOv2 uses anchor boxes and results better in detecting small objects. Dong et al. [18] used YOLOv2 to detect text in SVT and a self-made database. YOLOv3 is faster than YOLOv2 because of the complexity of the underlying architecture called Darknet. DetReco [19] detects objects as well as text using the structure of YOLOv3 and uses a state-of-the-art OCR to recognize objects with the class of text. Using the structure of YOLOv4 with a backbone of Cross Stage Partial Darknet53 (CSPD) and a 5-

parameter bounding box representation, R-Yolo [20] detects text in natural scene images with arbitrary rotation. As using the 5-parameter method can lead to the need for regression adjustment for some parameters, [21] proposes a new multi-directional bounding box representation along with an improved structure of YOLOv3 to detect text areas. Also, the CIOU loss is used as the base error and a step-by-step version of NMS to reduce calculations. The YOLO series have been used widely in object detection tasks

The most recent version, YOLOv5, is exceedingly smaller while maintaining a fast speed. To this end, this paper will take advantage of the high accuracy and speed of YOLOv5 and train a model to detect Farsi Text in images and videos.

III. PROPOSED METHOD

This study is done in two phases. First, a module based on the YOLOv5 model is proposed to detect Farsi text. Then in the rest of the research development of this framework by text recognition will be continued. Moreover, to prepare a proper benchmark for training and evaluating our method in the Farsi language, a new dataset of news videos is collected and arranged.

A. Model structure

As previously mentioned, YOLO is a one-stage object detection algorithm that requires much less memory and works about 1000 times faster than R-CNN and 100 times faster than Fast R-CNN. YOLO defines object detection as a regression problem. A convolutional neural network is once applied to the entire image and multiple bounding boxes with class probabilities are predicted simultaneously.

YOLO Version2 (YOLOv2) [22] is an improved version of YOLO which is faster and can work with different sizes. YOLOv3 [23] is an advanced version of the YOLO architecture. It better detects small objects compared to its predecessor YOLOv2 with the main feature of three layers at the output, each of which is designed to detect objects of different sizes. YOLOv4 [24] improved the previous versions in speed and gained an increase by 10% in mean average precision (mAP). The latest version of the YOLO series is YOLOv5, released in May 2020. Unlike the previous versions, YOLOv5 is implemented in PyTorch and not Darknet. A detailed diagram of the YOLOv5 architecture is shown in Fig. 1 This version outperforms all the previous versions and got near to EfficientDet [25] AP (average precision) with higher FPS (frames per second). The YOLOv5 model architecture consists of 3 main parts: model backbone, model neck, and model head.

Model backbone is mainly used as a feature extractor to extract important features of the input image. YOLOv5 uses CSPDarknet [24] as a feature extractor. CSP network has shown significant improvement in processing time with deeper networks. After the features are extracted, they are passed to the model neck to generate feature pyramids. Feature pyramids help the model to identify the same object at different scales.

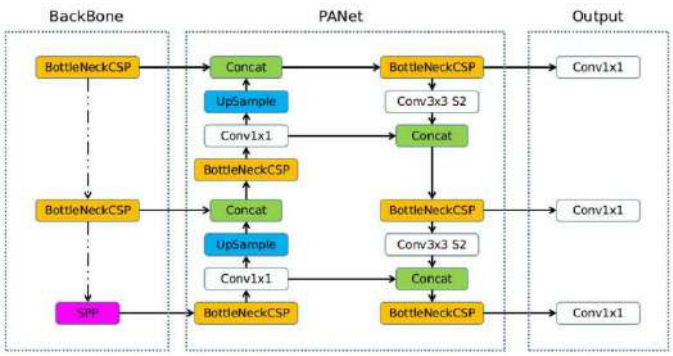


Figure 1. An overview of YOLO5

PANet [26] is used as the model neck in YOLOv5. The third part of the architecture is the model head and its duty is detection. YOLOv5 uses the same head as YOLOv3 and YOLOv4. It applies anchor boxes on features and generates final output vectors with class probabilities, objectness scores, and bounding boxes.

YOLOv5 is proposed in 4 versions. YOLOv5-s, YOLOv5-m, YOLOv5-l, and YOLOv5-x. These versions are distinguished from each other by the number of convolutional layers and subsamples used in the CSP block. Among the versions, there is a tradeoff between speed and accuracy. The small version is the fastest while the x version has the highest AP. YOLOv5 has an extensive market size and due to this, it can easily be adapted to text detection tasks. This paper used YOLOv5-l as it has been shown to have the highest accuracy according to the dataset size.

B. Dataset

To train and test YOLOv5 for detecting Farsi text in video frames, a dataset is needed. To the knowledge of the author, Unlike English and Chinese languages [27] [28], there is no dataset of images containing Farsi text lines in the input format of YOLOv5. Therefore, a dataset of news videos that are easily available and contain rich text areas with different styles and sizes, called FaMo Farsi Dataset, is generated. 95 news videos were gathered and 42 of them were selected for annotations. The selection of videos was based on the amount of text, different fonts and sizes, content genres, and of course the resolution of the video. The selected videos contain caption text, scene text, and variations in lighting and background. The total number of 21000 frames were annotated resulting in a text file and an image for each. The annotation was done with 2 classes; text and number. The number class was considered to help improve these systems in future work. Fig. 2 shows a few frames of the dataset.

C. Implementation details and Parameters

YOLOv5 offers SGD and ADAM optimizer and SGD is used in this research. The loss function of this model is also a version of Binary Cross-Entropy (BCE) implemented in Pytorch with the name of BCE With Logits Loss. This loss function uses BCE along with a sigmoid function in one single class which is more numerically stable than using a simple sigmoid function and then calculating BCE loss.

Figure 2. Samples of dataset frames

Epoch number was considered as 4000 (2000 * number of classes) but the training process could be stopped earlier if the model converges. The batch size of 64 could be suitable for this amount of data but due to memory limitations, it was set to 16. Initial weights of the network were selected randomly and the data was cached into memory for faster training. The learning rate had the initial value of 3.3298e-3 and the value increased to 0.01 for the SGD optimizer and started decreasing afterward along the training process. The changes in the learning rate are illustrated in Fig. 3.

Neural network training was done incrementally on a cloud base virtual machine called Google Colaboratory which is based on the Jupyter notebook. This virtual machine has the following specifications: 1xsingle core hyper-threaded Xeon Processors CPU @2.3Ghz, 1xTesla K80 GPU, 12.6 GB RAM, and 30 GB Disk memory. The training process ran on the GPU with 700 epochs and took about 150 hours. Due to a stable loss for the last few epochs, the training process was interrupted and no longer continued to reach 4000 epochs.

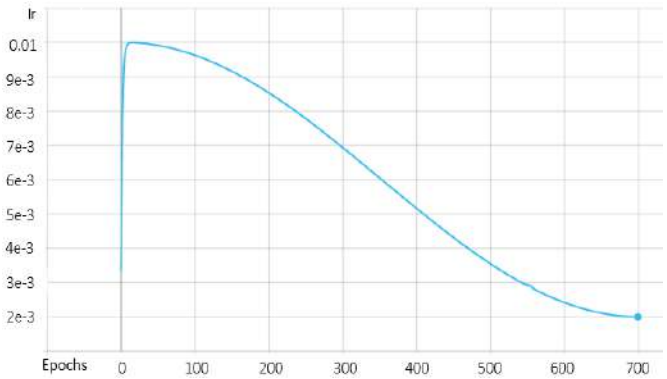


Figure 3. Changes of learning rate

IV. ANALISIS AND CONCLUSION

To evaluate our model, some experiments were done and results are analyzed based on standard evaluation metrics of this domain.

A. Dataset setup

After the video frames are labeled, the dataset is divided to train (66%), validation (16.5%) and test (16.5%) sets. The validation set is considered to prevent overfitting and help improve generalized detection. As the sets mentioned above were selected randomly, similar frames could have been placed in different sets. Therefore a few new videos from the same sources were chosen to evaluate the model more accurately. As the video frames in the dataset had different resolutions, they were first resized to 416 by 416 and then passed to the network.

B. Evaluation metrics

The Intersection over Union (IoU) is the area of overlap over the area of union between the ground truth and the predicted bounding box. In object detection, IoU is considered exceedingly important as it is used in calculating almost all evaluation metrics. To evaluate the detection model performance, mAP 0.5:0.95(mean average precision) is used and it corresponds to the average mAP with thresholds from 0.5 to 0.95 with a step size of 0.05. To observe the accuracy and purity of the model, precision and recall were calculated.

C. Experimental results

The trained model provides the user with information on the probability that a selected area of the image contains Farsi text. The application also builds a bounding box around the recognized text with information about the class.

Fig. 4 and Fig. 5 show the test results based on our dataset. As can be seen in Fig. 4, the model works pretty well in detecting fixed text with multi directions, colors, sizes, and styles. Moving text in video frames can have low quality and this might cause missing text areas in prediction.

Fig. 5(b) is an example of moving captions with low quality that are not detected as text. Another problem with the detection is that most videos containing subtitles or captions have a specific format (3 lines below each other with specific background colors). If by any chance, a line is transitioning or missing, the model might still detect the background area as text, Fig. 5(a). This problem could be fixed by expanding the dataset and considering different arrangements of text.

As can be seen in Fig. 6(c) the mAP metric has converged and increased significantly during the training process and at the end of 700 epochs, it has reached the value of 0.8017. The three other models of YOLOv5 were also trained on the same dataset and comparing the results in Table I. proofs that YOLOv5-l has better performance due to the large volume of the dataset. However, during this process, there were distortions due to the small batch size of 16. Increasing the batch size will help to decrease these distortions.

Figure 4. Positive test results based on FaMo dataset

(a) (b)

Figure 5. Negative results based on FaMo dataset. (a) False-positive.

(b) False Negative

Previous researches about text detection on the Farsi language were based on traditional methods. To compare the proposed model with similar work, recall and precision rates were used. According to Table II YOLOv5-l has the precision and recall rate of 97.75 and 94.46 respectively. These results show that YOLOv5-l outperforms other previous methods evaluated on similar datasets. This improvement is because of applying deep learning methods to text detection. The use of deep learning not only increases the robustness and efficiency of the model but also speeds up the training process.

Various loss functions are also computed along with training. One of the most important loss functions is box loss. Box-loss represents how fit the predicted bounding box is to the ground truth bounding box. The smaller the box-loss, the less none-text area detected. This means that bounding boxes will fit text lines more accurately.

The labeling of the dataset was done with horizontal rectangles and the predicted bounding boxes are also in this format. This will decrease the box loss from 0.1081 to 0.0191 during training. However, this does not mean that the predicted bounding boxes are well fitted to the text areas. The bounding box around diagonal text lines will include background areas.

This does not reduce performance much but using a multi-directional detection model will predict more robust bounding boxes.

TABLE I. COMPARISON OF DIFFERENT YOLOV5 MODELS ON THE SAME DATASET ACCORDING TO THE MAP METRIC

Type of model YOLOv5	mAP 0.5-0.95
YOLOv5-s	0.7848
YOLOv5-m	0.7884
YOLOv5-l	0.8017
YOLOv5-x	0.7972

TABLE II. COMPARISON BETWEEN DIFFERENT METHODS FOR FARSI/ARABIC TEXT DETECTION WITH SIMILAR DATASETS

Approach	Precision rate	Recall rate
corner detection [4]	72.80	87.54
cdge map + color properties [6]	80.80	86.50
corner detection + texture intensity picture [5]	91.38	87.22
proposed approach	97.75	94.46

D. Conclusion

This paper mainly introduces the training of target-detection framework YOLOv5-l applied to Farsi text detection in news images and videos. The model was built based on a self-made database containing news videos. The proposed method achieved an overall mAP of 0.9736 and the resulting test images show high accuracy in detecting text lines with complex backgrounds and varying sizes.

The framework will be applicable to video texts like signboards, subtitles from movies, natural, and wild scene texts with different fonts.

E. Future work

Although this research is a novel attempt in Farsi text detection, its essential improvement is undeniable. Possible future research directions can include:

- Improving the dataset: expanding dataset by street images, handwritten images, and a wide variety of number text.
- Recognition: Feeding the detected text areas to an OCR engine to perform the recognition task.

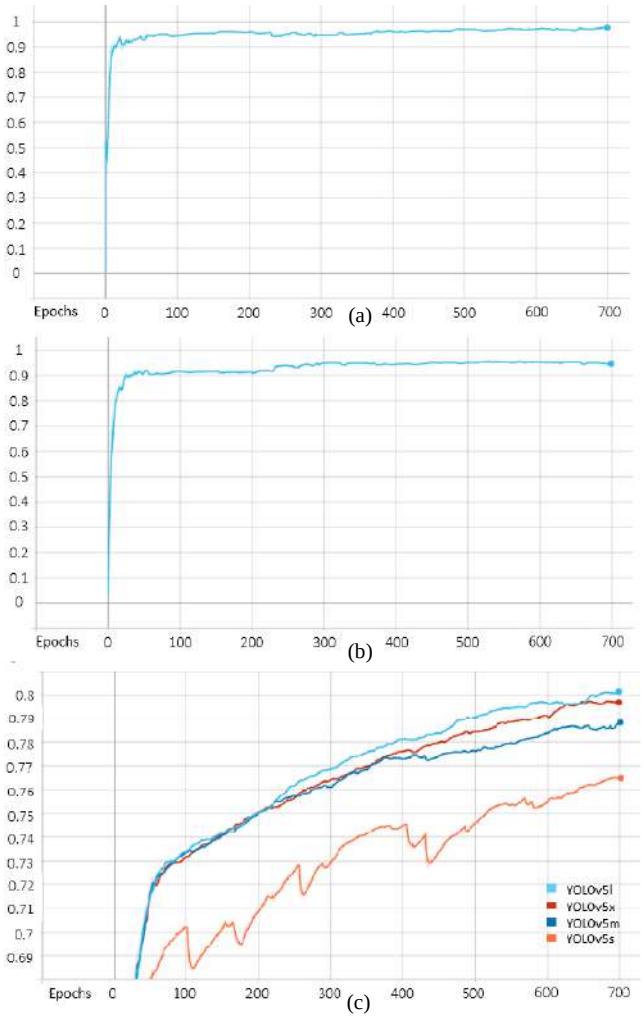


Figure 6. Metric values in the learning process of YOLOv5 model. (a) Precision. (b) Recall. (c) mAP 0.5:0.95, comparing all 4 models

REFERENCES

[1] D. Snegireva and A. Perkova, "Traffic Sign Recognition Application Using Yolov5 Architecture," Proceedings - 2021 International Russian Automation Conference, RusAutoCon 2021, pp. 1002-1007, 2021.

[2] A. Ali and M. A. Ali, "Blind navigation system for visually impaired using windowing-based mean on Microsoft Kinect camera," 2017 Fourth International Conference on Advances in Biomedical Engineering (ICABME), pp. 1-4, 2017.

[3] G. Praveen and R. Paily, "Blind Navigation Assistance for Visually Impaired Based on Local Depth Hypothesis from a Single Image," Procedia Engineering, vol. 64, 2013.

[4] M. Moradi, S. Mozaffari and A. A. Orouji, "Farsi/Arabic text extraction from video images by corner detection," 2010 6th Iranian Conference on Machine Vision and Image Processing, pp. 1-6, 2010.

[5] M. Moradi and S. Mozaffari, "Hybrid approach for Farsi/Arabic text detection and localisation in video frames," IET Image Processing, vol. 7, no. 2, pp. 154-164, 2013.

[6] M. Darab and M. Rahmati, "A hybrid approach to localize farsi text in natural scene images," Procedia Computer Science, vol. 13, pp. 171-184, 2012.

[7] B. D. L, "DRT: A Unified Framework for Text Detection, Recognition and Tracking in Video; DRT: A Unified Framework for Text Detection, Recognition and Tracking in Video," 2015.

[8] M. M. Tabari, J. A. Aghamaleki and A. Ghorbani, " تشخیص متون فارسی در تصاویر بر پایه توصیفگر SFIT و لغت نامه فازی [Farsi text detection in images based on SIFT descriptor and fuzzy dictionary]," 3rd International Conference on Applied Research in Computer Engineering and Information Technology, Tehran, 2016.

[9] H. Chen, S. S. Tsai, G. Schroth, D. M. Chen, R. Grzeszczuk and B. Girod, "Robust text detection in natural images with edge-enhanced Maximally Stable Extremal Regions," 2011 18th IEEE International Conference on Image Processing, pp. 2609-2612, 2011.

[10] X.-C. Yin, X. Yin, K. Huang and H.-W. Hao, "Robust Text Detection in Natural Scene Images," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 36, pp. 970-983, 2014.

[11] D. Cao, Y. Zhong, L. Wang, Y. He and J. Dang, "Scene Text Detection in Natural Images: A Review," Symmetry, vol. 12, p. 1956, 2020.

[12] H. Li and H. Lu, "At-text: Assembling text components for efficient dense scene text detection," Future Internet, vol. 12, no. 11, pp. 1-14, 2020.

[13] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," IEEE Transactions on Pattern Analysis & Machine Intelligence, vol. 39, pp. 1137-1149, 2017.

[14] R. B. Girshick, "Fast R-CNN," 2015 IEEE International Conference on Computer Vision (ICCV), pp. 1440-1448, 2015.

[15] Y. Jiang et al., "R2 CNN: Rotational Region CNN for Arbitrarily-Oriented Scene Text Detection," 2018 24th International Conference on Pattern Recognition (ICPR), pp. 3610-3615, 2018.

[16] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu and A. C. Berg, "SSD: Single Shot MultiBox Detector," Lecture Notes in Computer Science, pp. 21-37, 2016.

[17] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779-788, 2016.

[18] D. Haifeng and H. Siqi, "Natural scene text detection based on {YOLO} V2 network model," in Journal of Physics: Conference Series, 2020.

[19] F. Zhang, J. Luan, Z. Xu and W. Chen, "DetReco: Object-Text Detection and Recognition Based on Deep Neural Network," Mathematical Problems in Engineering, 2020.

[20] X. Wang, S. Zheng, C. Zhang, R. Li and L. Gui, "R-YOLO: A Real-Time Text Detector for Natural Scenes with Arbitrary Rotation," Sensors, vol. 21, p. 888, 2021.

[21] L. Xiao, P. Zhou, K. Xu and X. Zhao, "Multi-Directional Scene Text Detection Based on Improved YOLOv3," Sensors, vol. 21, 2021.

[22] X. Huang et al., "PP-YOLOv2: A Practical Object Detector," ArXiv, vol. abs/2104.10419, ArXiv.

[23] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," ArXiv, vol. abs/1804.02767, 2018.

[24] A. Bochkovskiy, H.-Y. M. Liao and C.-Y. Wang, "YOLOv4: Optimal Speed and Accuracy of Object Detection," ArXiv, vol. abs/2004.10934, 2020.

[25] M. Tan, R. Pang and Q. V. L. V. Le, "EfficientDet: Scalable and Efficient Object Detection," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 10778-10787, 2020.

[26] K. Wang, J. H. L. Liew, Y. Zou, D. Zhou and J. Feng, "PANet: Few-Shot Image Semantic Segmentation with Prototype Alignment," 2019 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 9196-9205, 2019.

[27] "MSRA-TD500 (MSRA Text Detection 500 Database)," [Online]. Available: https://paperswithcode.com/dataset/msra-td500.

[28] "COCO-Text," [Online]. Available: https://paperswithcode.com/dataset/coco-text.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)
March 2-4, 2022, Higher Education Complex of Bam

Feature selection for multi-label text data: An ensemble approach using geometric mean aggregation

Mohsen Miri[1], Mohammad Bagher Dowlatshahi[2], Amin Hashemi[3]
Department of Computer Engineering, Faculty of Engineering, Lorestan University, Khorramabad, Iran
[1] miri.mo@fe.lu.ac.ir
[2] dowlatshahi.mb@lu.ac.ir
[3] hashemi.am@fe.lu.ac.ir

***Abstract*—Text datasets have many terms, which decrease the classification accuracy. According to the high-dimensional text data, there are more challenges for these methods. Each classification method has strengths and weaknesses in its feature selection function. Therefore, ensembling should be used for better classifications and exploitation of strengths. In this paper, for the first time, we have presented an ensemble multi-label (ML) feature selection method for the text datasets using the Geometric-Mean (GM) aggregation approach. For this purpose, we have used four multi-label feature selection (MLFS) algorithms with different structures to achieve a good result. Then, the performance and results obtained by the GM method are compared with the four algorithms and based on the six classification criteria on three ML datasets with text domains. According to the obtained outputs, it is possible to realize the ability of the GMA (proposed method) method in using strengths and ignoring weaknesses in the path of feature selection, and be more accurate. Finally, according to experiments and obtained results, superiority of the GMA over other methods can be well seen.**

***Index Terms* —Ensemble learning; Multi-label feature selection; Text classification; Geom-mean aggregation**

I. INTRODUCTION

Text data can be placed in different categories simultaneously, and text classification is used to identify different categories of text [1], [2]. Usually, some redundant and inessential features in dataset negatively affect or not affect data analysis output [3], which to remove these features, we can use Feature Selection (FS). FS can increase classification accuracy and reduce data dimensions, computational cost, learning time, and storage complexity [4], [5].

The FS methods can be divided into three categories of the filters, wrappers, and embedded [6]–[9]. The wrapper methods are more accurate than the filter and embedded methods, and the filter-based methods are faster. In addition to the above categories, FS has three modes based on the interaction with class labels: supervised, semi-supervised, and unsupervised [6], [10].

Rankings obtained by different algorithms have different accuracies on specific data. In this regard, to improve the performance of FS methods on ML text data, an ensemble FS approach is applied for the first time. Initially, the FS operations are applied to the ML data. At this stage, the rankings of four methods are combined based on GM, and the final order of the features will be obtained. This procedure ignores the weaknesses of algorithms and selects features based on the strengths of FS rankings.

The continuation of the article is as follows: Section II, will introduce some related works, ML learning, and ensemble FS. In Section III, the details of the GMA method are describe. In Section IV, experiment and performance analysis of the proposed method and comparisons are performed. Finally, the article ended with a discussion and conclusion.

II. RELATED WORKS AND FUNDAMENTAL CONCEPT

A. Related works

Some MLFS methods are briefly introduced below, which are used in the structure of the GMA method. In [11] introduced a FS algorithm called MCLS. These methods use multi-pupose learning and transfer the logical space to the Euclidean space. Also, supervision information and local data features affect the final selection. Paniri et al. [12] presented a MLFS based on Ant Colony Optimization, called MLACO. The maximum and minimum relationships between features and tags are considered based on supervised and unsupervised exploratory performances to calculate the redundancy between features.

In [13], another FS algorithm for ML data called MDFS was introduced. Zhang et al. used manifold regularization to produce the low dimensional embedding from the main attributes space.

In this algorithm considering L2-norm regularization as a learning framework, the FS was enforced to search for discriminative attributes. One of the main problems in FS is to ignore the effect of interaction information on the final choice. To solve this problem, Zhang et al. [14], considering the impact of interaction information on calculating the importance of features, introduced a FS algorithm called LRFS.

B. Multi-label learning

Machine learning is implemented on two categories of single-label and ML data [8], [15]–[18]. ML datasets (Fig.1) have two or more labels. In this data, each instance has a vector of features as $X_i=(x_{i1},x_{i2},x_{i3}, \ldots,x_{iM})$ and also the labels as $Y_i= (y_{i1},y_{i2},y_{i3}, \ldots,y_{iL})$. M show the number of features and L show the number of labels. In ML data, each label has a binary value. ML learning uses the N instances of data, known as training data, to make a model that can predict the label for new samples [7], [19].

C. Ensemble feature selection

Dietterich [20] presented ensemble FS[21], [22] based on the aggregation feature selector. The ensemble techniques can increase accuracy and improve the Robustness of FS methods. The ensemble method tries to obtain the overall optimal rankings by combining the optimal local outputs of the feature selectors [23]. According to the type of electors, ensemble methods are divided into two categories [24]. It is homogeneous when all selectors are of the same type; otherwise, it is heterogeneous (Fig.2). See [24] for more details.

III. PROPOSED METHOD

In this section, we describe our proposed algorithm in detail. GMA is a heterogeneous ensemble method for text data that uses four MLFS algorithms. Each FS algorithm gets its ranking. We put the obtained ranks in a matrix. Finally, we get the final rank by aggregating the rankings using the GM. Fig.3 shows a summary of the proposed method graphically.

X_1	X_1	⋯	X_M	Y_1	Y_2	⋯	Y_L
X_{11}	X_{12}	⋯	X_{1M}	0	1	⋯	0
X_{21}	X_{22}	⋯	X_{2M}	1	0	⋯	0
⋮	⋮	⋱	⋮	⋮	⋮	⋱	⋮
X_{N1}	X_{N2}	⋯	X_{NM}	0	1	⋯	1

Figure 1. ML data structure

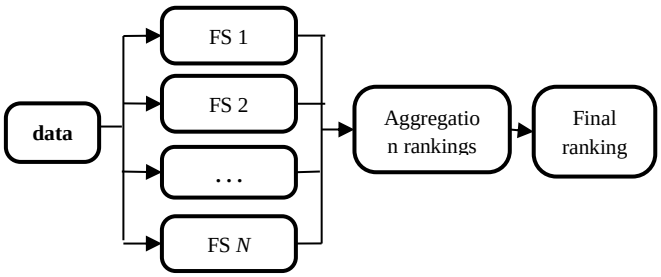


Figure 2. Heterogeneous ensemble ranking

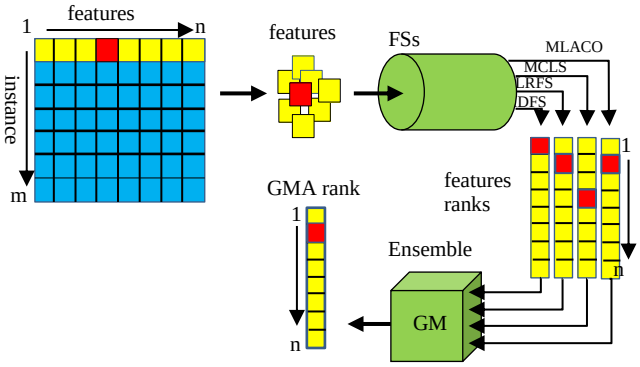


Figure 3. Summary of the GMA

As you can see in Fig.3, the ensemble algorithm can be considered as a voting system. FS methods (MLACO, LRFS, MCLS, MDFS) are the voters, and the features are the candidates. Each voter sends its votes (feature ranks) to the system, and then the votes are calculated based on GM.

Now we explain this method step by step according to Algorithm 1. We introduce an empty matrix $N \times 1$ (AY) for the final ranking in the first step. In the next four steps, we execute the four MLFS algorithms: MDFS, MCLS, LRFS, and MLACO on the text datasets and place the obtained rankings in $R1$ to $R4$ vectors. We combine the rankings of these four methods to achieve better ranks and more important features to reduce the impact of the redundant features on the rankings.

$$R(j)=\{r_{j,1},r_{j,2},\ldots\ldots,r_{j,M}\},\ j\in[1,4]algorithms$$

In step 6, we sort the obtained rankings, and in step 7, we place them in the AR matrix. Each row represents the obtained ranks of a feature in four different algorithms. Also, each column in this matrix represents the ranking of features by an algorithm.

Since the AR matrix is based on the ranking of attributes and we are looking to maximize the importance, we need to normalize this matrix. After normalizing and scoring the ranks in the range of [0,1], any rank with a lower score (close to 0) is more important, and also, the lowest importance is related to the highest score (maximum score equal to 1). In step 8, we perform the normalization of the AR matrix using the following equation.

$$AR=\begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ r_{21} & r_{22} & r_{23} & r_{24} \\ \vdots & \vdots & \vdots & \vdots \\ r_{M1} & r_{M2} & r_{M3} & r_{M4} \end{bmatrix}$$

$$NAR=\frac{AR}{M} \quad (1)$$

In step 9, the AR matrix is sent to the geometric-mean function to aggregate the ranks according to equations 2 and 3 related to GM. Finally, the AY vector represents the final ranking obtained based on the GMA method. The proposed algorithm is then compared with the other four algorithms used, discussed in the next session.

$$X(i)=\frac{1}{n}\sum_{j=1}^{n} \log NAR(i,j) \quad (2)$$

$$AY(i)=e^{X(i)}; \quad (3)$$

$$AY=[r_1,r_2, r_3,r_4, \ldots., r_M]$$

I. EXPERIMENTS

Since the ensemble FS strategy has not been used on ML data so far to measure the performance of our method, we compare GMA with the four basic MLFS algorithms discussed in this article.

A. Datasets

We used three text datasets to demonstrate the superiority of our method over other algorithms. Actual ML text data are selected from the most widely used text mining datasets. They have different samples, features, and labels in terms of number. The descriptions of these datasets can be seen in Table1.

B. Evaluation metrics

Suppose $T = \{(x_i, Y_i): i = 1,\ldots,n\}$ be a test set, and $L=\{l_j:j=1,\ldots,q\}$ be the set of all labels where $Y_i \subseteq L$, and h be the classifier. According to sample x_i, the labels predicted by the classifier are denoted as Z_i, and $f()$ is a function that is returned by multiple learning systems. We use the following criteria to measure the performance of the GMA method and the other four algorithms and compare them [16], [19]:

1) *Accuracy:*

The degree of agreement between the measured value and the actual value.

$$Accuracy(h,T)=\frac{1}{n}\sum_{i=1}^{n}\frac{|Y_i \Delta Z_i|}{|L|} \quad (4)$$

2) *Hamming loss:*

The fraction of labels that are incorrectly predicted.

$$Hamming\_loss(h,T)=\frac{1}{n}\sum_{i=1}^{n}\frac{|Y_i \Delta Z_i|}{|L|}, \quad (5)$$

where Δ is the symmetric difference between two sets.

3) *One error:*

one error is the rejection of a true null hypothesis. It measures how many times the ranked label is unrelated to the instance labels.

$$One\_error(f)=\frac{1}{n}\sum_{i=1}^{n}\left[\left[\arg\ \underset{y\in Y_i}{\arg\ \max} f(x_i,y)\right]\right]\notin Y_i \quad (6)$$

4) *Coverage:*

Measures the average number of steps required to cover all relevant labels.

$$Coverage(f)=\frac{1}{n}\sum_{i=1}^{n}\max_{y\in Y_i} rank_f(x_i,y)-1 \quad (7)$$

5) *Ranking loss:*

The objective of Ranking Losses is to predict relative distances between inputs.

$$Rank\_loss(f)=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{|Y_i||\bar{Y}_i|}\left|\{(y', y)|f(x_i,y') \leqslant f(x,y),(y',y'') \in Y_i \times \bar{Y}_i\}\right| \quad (8)$$

6) *Average precision:*

A measure that combines Recall and Precision measures for ranked retrieval results.

$$Ave\_pre(f)=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{|Y_i|}\sum_{y\in Y_i}\frac{\{y'|\ rank_f(x,y')\leqslant rank_f(x_i,y),y'\in Y_i\}}{rank_f(x_i,y)} \quad (9)$$

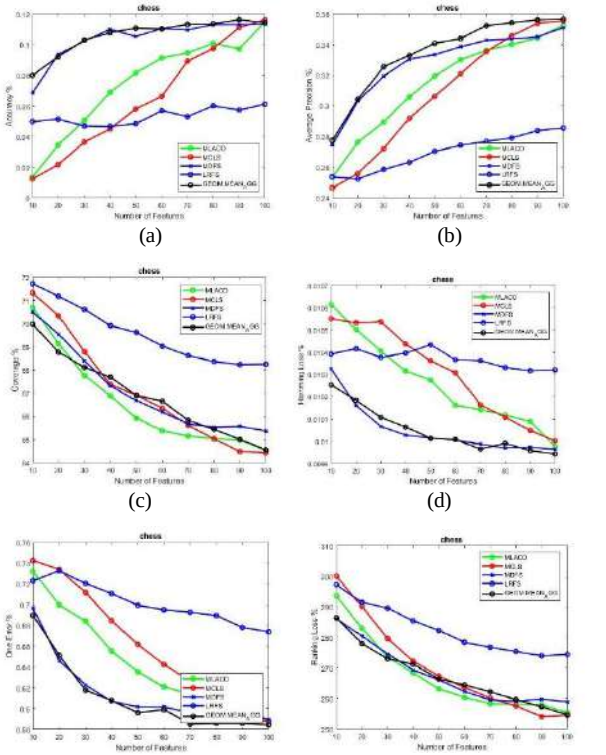
C. Result

To measure the performance of the GMA against the four comparing algorithms, the multi-label K-nearest neighbor (KNN) classifier [25] with ten neighbors is used, which is compared based on the output of this classifier. At the beginning of the experiments, the data are divided between the training and test sets with 60% and 40% ratios.

Ten equal validations have been used to evaluate this performance of algorithms, where m features are selected as representative from the features extracted from each algorithm, and the test is performed on them. The number of representatives is between 10 and 100 and a factor of 10 ($m \in \{10, 20, 30, 40, 50, 60, 70, 80, 90, 100\}$). All results are calculated based on the average of ten separate runs of the algorithms in different datasets.

Table 1. Dataset descriptions

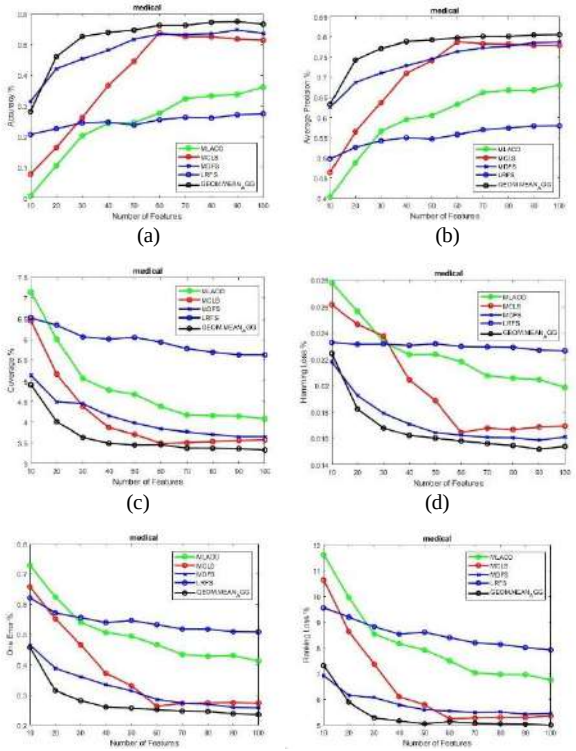
Num	*Dataset*	*Domain*	*Samples*	*Features*	*Labels*
1	Chess	Text	1675	585	227
3	Medical	Text	978	1449	45
4	Philosophy	Text	3971	842	233



(a) (b) (c) (d)

(e) (f)

Figure 4. Outputs for the Chess dataset based on: (a) Accuracy, (b) Ave-precision, (c) Coverage (d) Hamming loss (e) One-E (f) Ranking loss

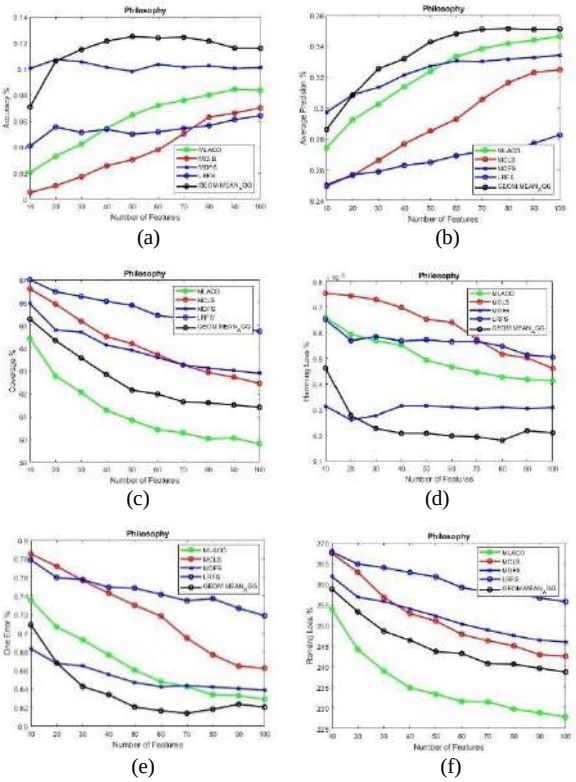


(a) (b)

(c) (d)

(e) (f)

Figure 5. Outputs for the Medical dataset based on: (a) Accuracy, (b) Ave-precision, (c) Coverage (d) Hamming loss (e) One-E (f) Ranking loss



(a) (b)

(c) (d)

(e) (f)

Figure 6. Outputs for the Philosophy dataset based on: (a) Accuracy, (b) Ave-precision, (c) Coverage (d) Hamming loss (e) One-E (f) Ranking loss

The results obtained for the GMA method on different datasets are compared with the four comparing algorithms, are shown in Fig.4 to Fig.6. In the diagrams drawn in these figures, the vertical axis shows the value of the desired performance criterion, and the horizontal axis indicates the number of selected features. As can be seen, the proposed algorithm performs better than the other four methods based on different criteria. According to the results obtained in Figs. 4-6, the GMA is completely superior to other methods used. In this way, based on the criteria of accuracy, ave-precision, hamming loss, and one-error, the superiority of the results related to the GM-aggregation method can be clearly seen. However, this method has not achieved complete superiority over competitors based on the coverage, and ranking loss criteria. GM-Aggregation based on coverage and ranking loss in philosophy data is relatively weaker than MLACO but more robust than the other three methods. This action explicitly demonstrates the desired reality of the problem, which is to take advantage of the strengths and ignore the weaknesses of other algorithms.

Tables 2 to 6 also present the average of the total values obtained for each run on different datasets.

Table 2. Average Hamming loss of 10 final subsets of features

Algorithm	*Chess*	*Philosophy*	*Medical*
MLACO	0.0103	0.0095	0.0225
LRFS	0.0104	0.0096	0.0230
MCLS	0.0103	0.0096	0.0198
MDFS	**0.0100**	0.0093	0.0173
GM_Agg	**0.0100**	**0.0092**	**0.0167**

Table 3. Mean Ave-precision of 10 final subsets of features

Algorithm	*Chess*	*Philosophy*	*Medical*
MLACO	0.3149	0.3210	0.5968
LRFS	0.2698	0.2665	0.5521
MCLS	0.3085	0.2896	0.7028
MDFS	0.3283	0.3226	0.7383
GM_Agg	**0.3345**	**0.3347**	**0.7737**

Table 4. Average Coverage of 10 final subsets of features

Algorithm	*Chess*	*Philosophy*	*Medical*
MLACO	**66.5455**	**61.1999**	4.8581
LRFS	69.5518	65.7349	5.9646
MCLS	67.0596	64.1540	4.1229
MDFS	67.0731	63.9414	4.0812
GM_Agg	66.8886	62.6343	**3.6376**

Table 5. Average One-error of 10 final subsets of features

Algorithm	*Chess*	*Philosophy*	*Medical*
MLACO	0.6448	0.6658	0.5065
LRFS	0.7016	0.7453	0.5421
MCLS	0.6584	0.7204	0.3736
MDFS	0.6154	0.6526	0.3206
GM_Agg	**0.6103**	**0.6367**	**0.2784**

Table 6. Average Ranking loss of 10 final subsets of features

Algorithm	*Chess*	*Philosophy*	*Medical*
MLACO	267.1712	**235.4928**	8.1439
LRFS	282.5100	260.9340	8.5389
MCLS	269.9056	251.6043	6.5073
MDFS	267.6054	252.0458	5.8018
GM_Agg	**267.3075**	245.3819	**5.4051**

Table 7. Average Accuracy of 10 final subsets of features

Algorithm	*Chess*	*Philosophy*	*Medical*
MLACO	0.0748	0.0610	0.2435
LRFS	0.0531	0.0539	0.2487
MCLS	0.0654	0.0375	0.3938
MDFS	0.1041	0.1022	0.4875
GM_Agg	**0.1062**	**0.1140**	**0.5196**

D. Discussion

As seen in the previous step, the performance of the proposed method was evaluated and compared with four other algorithms based on six criteria. The result was that the GMA algorithm had the best performance compared to competitors in four important criteria and won this comparison. The performance of the method based on the coverage, and ranking loss criteria is also acceptable because it is better than other in Medical dataset, and is competes to win with other methods in chess datasets. Based on what has been said and seen, it is possible to reach the important point of the problem, which the GMA algorithm uses the strengths of the other four algorithms to obtain the ranking. One of the ensemble technique characteristics is to take advantage of this issue, and the other is to ignore the weaknesses of algorithms. The proposed algorithm using the GM-aggregation method has taken good advantage of these two advantages.

II. CONCLUSION

This paper presents an ensemble MLFS method using geometric mean aggregation for text data for the first time. Four MLFS algorithms MLACO, MCLS, LRFS, and MDFS, are used. In this method, we build our model with 60% of the data. In this way, we first implement ML algorithms on the data so that each algorithm presents its ranking. We put the obtained ranks in a whole matrix. Each row of this matrix corresponds to the ranks obtained for each attribute. Furthermore, the ranking of the properties obtained by each method is shown in the columns. In the following, using the geometric mean method, we combine the rankings obtained by different methods. Generally, with the experiments performed on the proposed method and comparison with the other four algorithms, the superiority of the GMA method based on Hamming Loss, One error, Accuracy, Average Precision, and relative superiority based on Coverage and Ranking Loss is clearly shown.

ACKNOWLEDGMENT

This research has not received any special privileges and financial and non-financial assistance from various organizations.

REFERENCES

[1] J. Lee, I. Yu, J. Park, and D. W. Kim, "Memetic feature selection for multilabel text categorization using label frequency difference," *Inf. Sci. (Ny).*, vol. 485, pp. 263–280, Jun. 2019, doi: 10.1016/j.ins.2019.02.021.

[2] D. S. Guru, M. Suhil, S. K. Pavithra, and G. R. Priya, "Ensemble of Feature Selection Methods for Text Classification: An Analytical Study," 2018, pp. 337–349.

[3] M. B. Dowlatshahi, V. Derhami, and H. Nezamabadi-pour, "A novel three-stage filter-wrapper framework for miRNA subset selection in cancer classification," *Informatics*, vol. 5, no. 1, 2018, doi: 10.3390/informatics5010013.

[4] J. Miao and L. Niu, "A Survey on Feature Selection," *Procedia Comput. Sci.*, vol. 91, pp. 919–926, Jan. 2016, doi: 10.1016/J.PROCS.2016.07.111.

[5] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-

Pour, "Gravitational Search Algorithm: Theory, Literature Review, and Applications," *Handb. AI-based Metaheuristics*, pp. 119–150, 2021.

[6] A. Hashemi, M. Bagher Dowlatshahi, and H. Nezamabadi-pour, "A pareto-based ensemble of feature selection algorithms," *Expert Syst. Appl.*, vol. 180, p. 115130, Oct. 2021, doi: 10.1016/j.eswa.2021.115130.

[7] A. Hashemi, M. Bagher Dowlatshahi, and H. Nezamabadi-pour, "An efficient Pareto-based feature selection algorithm for multi-label classification," *Inf. Sci. (Ny).*, vol. 581, pp. 428–447, Dec. 2021, doi: 10.1016/j.ins.2021.09.052.

[8] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-pour, "MGFS: A multi-label graph-based feature selection algorithm via PageRank centrality," *Expert Syst. Appl.*, vol. 142, p. 113024, 2020.

[9] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-pour, "VMFS: A VIKOR-based multi-target feature selection," *Expert Syst. Appl.*, vol. 182, p. 115224, Nov. 2021, doi: 10.1016/j.eswa.2021.115224.

[10] B. Venkatesh and J. Anuradha, "A review of Feature Selection and its methods," *Cybern. Inf. Technol.*, vol. 19, no. 1, pp. 3–26, 2019, doi: 10.2478/CAIT-2019-0001.

[11] R. Huang, W. Jiang, and G. Sun, "Manifold-based constraint Laplacian score for multi-label feature selection," *Pattern Recognit. Lett.*, vol. 112, pp. 346–352, Sep. 2018, doi: 10.1016/j.patrec.2018.08.021.

[12] M. Paniri, M. B. Dowlatshahi, and H. Nezamabadi-pour, "MLACO: A multi-label feature selection algorithm based on ant colony optimization," *Knowledge-Based Syst.*, vol. 192, p. 105285, Mar. 2020, doi: 10.1016/j.knosys.2019.105285.

[13] J. Zhang, Z. Luo, C. Li, C. Zhou, and S. Li, "Manifold regularized discriminative feature selection for multi-label learning," *Pattern Recognit.*, vol. 95, pp. 136–150, Nov. 2019, doi: 10.1016/j.patcog.2019.06.003.

[14] P. Zhang, G. Liu, and W. Gao, "Distinguishing two types of labels for multi-label feature selection," *Pattern Recognit.*, vol. 95, pp. 72–82, Nov. 2019, doi: 10.1016/j.patcog.2019.06.004.

[15] M. Paniri, M. B. Dowlatshahi, and H. Nezamabadi-pour, "Ant-TD: Ant colony optimization plus temporal difference reinforcement learning for multi-label feature selection," *Swarm Evol. Comput.*, vol. 64, p. 100892, Jul. 2021, doi: 10.1016/j.swevo.2021.100892.

[16] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-Pour, "A bipartite matching-based feature selection for multi-label learning," *Int. J. Mach. Learn. Cybern.*, vol. 12, no. 2, pp. 459–475, Feb. 2021, doi: 10.1007/s13042-020-01180-w.

[17] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-pour, "MFS-MCDM: Multi-label feature selection using multi-criteria decision making," *Knowledge-Based Syst.*, vol. 206, p. 106365, 2020, doi: https://doi.org/10.1016/j.knosys.2020.106365.

[18] A. Hashemi and M. B. Dowlatshahi, "MLCR: a fast multi-label feature selection method based on K-means and L2-norm," in *2020 25th International Computer Conference, Computer Society of Iran (CSICC)*, 2020, pp. 1–7.

[19] S. Kashef, H. Nezamabadi-pour, and B. Nikpour, "Multilabel feature selection: A comprehensive review and guiding experiments," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 8, p. e1240, 2018, doi: 10.1002/widm.1240.

[20] T. G. Dietterich, "Ensemble Methods in Machine Learning," 2000, pp. 1–15.

[21] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-Pour, "A bipartite matching-based feature selection for multi-label learning," *Int. J. Mach. Learn. Cybern.*, vol. 12, no. 2, pp. 459–475, 2021, doi: 10.1007/s13042-020-01180-w.

[22] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-pour, "Minimum redundancy maximum relevance ensemble feature selection: A bi-objective Pareto-based approach," *J. Soft Comput. Inf. Technol.*, 2021.

[23] B. Schowe and K. Morik, "Fast-Ensembles of Minimum Redundancy Feature Selection," 2011, pp. 75–95.

[24] V. Bolón-Canedo and A. Alonso-Betanzos, "Ensembles for feature selection: A review and future trends," *Inf. Fusion*, vol. 52, pp. 1–12, 2019, doi: 10.1016/j.inffus.2018.11.008.

[25] M.-L. Zhang and Z.-H. Zhou, "ML-KNN: A lazy learning approach to multi-label leaming," *Pattern Recognit.*, vol. 40, pp. 2038–2048, 2007, doi: 10.1016/j.patcog.2006.12.019.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Folding Theory Applied to Integral EQ-alegbras

1st Akbar Paad, 2nd Azam Jafari

1,2nd Department of Mathematics, *University of Bojnord*, Bojnord, Iran, akbar.paad@gmail.com,

***Abstract*—In this paper, In this paper, the notion of n-fold integral EQ-algebras and n-fold integral prefilters are introduced and several characteristics of them are presented. Moreover, relation among n-fold integral prefilters and some type of other prefilter such n-fold positive implicative and n-fold obstinate prefilters in EQ-algebras are studied. Finally, it is proved that a filter of a EQ-algebra L is an n-fold integral filter if and only if L/F is an n-fold integral EQ-algebra.**

***Index Terms*—EQ-algebra, n-fold integral prefilter, n-fold integral EQ-algebra.**

I. Introduction

recently, a special algebra called EQ-algebra has been introduced by Novák [7]. One of the motivations was to introduce a special algebra as the correspondence of truth values for high-order fuzzy type theory (FTT)[8] that generalizes the system of classical type theory [1] in which the sole basic connective is equality. Analogously, the basic connective in (FTT) should be fuzzy equality. Another motivation is from the equational style of proof in logic. It has three connectives: meet $\wedge$, product $\odot$ and fuzzy equality $\sim$. EQ-algebras are interesting and important for studying and researching and residuated lattices are particular cases of EQ-algebras. In fact, EQ-algebras generalize non-commutative residuated lattices [4]. From the point of view of logic, the main difference between residuated lattices and EQ-algebras lies in the way the implication operation is obtained. While in residuated lattices it is obtained from (strong) conjunction, in EQ-algebras it is obtained from equivalence. Consequently, the two kinds of algebras differ in several essential points despite their many similar or identical properties. The prefilter theory plays a fundamental role in the general development of EQ-algebras. From a logical point of view, various prefilters correspond to various sets of provable formulas. Some types of prefilters in EQ-algebras based on logical algebras have been widely studied [2, 5, 6] and some important results have been obtained. Since residuated lattices (BL-algebras, $MV-$algebras, MTL-algebras, R_0-algebras) are EQ-algebras, it is natural to extend some notions of residuated lattices to EQ-algebras and study some of their properties. The notions of integral BL-algebras, EQ-algebras and integral filters of them had defined by Borzooei and Paad [3] , [11] and some properties of them are studied. This motivates us to extend this notions to EQ algebras, so the notions n-fold integral prefilters and n-fold integral EQ-algebras are introduced.

II. Preliminaries

Definition 1. *[9] An EQ-algebra is an algebra $(L, \wedge, \odot, \sim, 1)$ of type $(2, 2, 2, 0)$ satisfying the following axioms:*
$(e1)$ $(L, \wedge, 1)$ *is a $\wedge$-semilattice with top element 1. We set $x \leq y$ if and only if $x \wedge y = x$,*
$(e2)$ $(L, \odot, 1)$ *is a commutative monoid and $\odot$ is isotone with respect to $\leq$,*
$(e3)$ $x \sim x = 1$,
$(e4)$ $((x \wedge y) \sim z) \odot (s \sim x) \leq z \sim (s \wedge y)$,
$(e5)$ $(x \sim y) \odot (s \sim t) \leq (x \sim s) \sim (y \sim t)$,
$(e6)$ $(x \wedge y \wedge z) \sim x \leq (x \wedge y) \sim x$,
$(e7)$ $x \odot y \leq x \sim y$,
For all $s, t, x, y, z \in L$.

For all $x, y \in L$, we put

$$x \to y = (x \wedge y) \sim x, \qquad \tilde{x} = x \sim 1.$$

The derived operation $\to$ is called implication and if L contains a bottom element 0, then we may define the unary operation $\neg$ on L by $\neg x = x \sim 0$.

Definition 2. *[4, 9] Let L be an EQ-algebra. Then we say that it is*
(i) *Separated, if $x \sim y = 1$ implies $x = y$, for all $x, y \in L$.*
(ii) *Spanned if it contains a bottom element 0 and $\tilde{0} = 0$.*
(iii) *Residuated, if $(x \odot y) \wedge z = x \odot y$ if and only if $x \wedge ((y \wedge z) \sim y) = x$, for all $x, y, z \in L$.*

Definition 3. *[2, 6, 13] Let L be an EQ-algebra and $\emptyset \neq F \subseteq L$. Then*
F is called a prefilter *of L if it satisfies for all $x, y \in L$,*
(i) $1 \in F$,
(ii) $x \in F$, $x \to y \in F$ *imply that* $y \in F$.
A prefilter F is said to be a filter *if it satisfies:*
If $x \to y \in F$, then $(x \odot z) \to (y \odot z) \in F$ $(*)$, for all

$x, y, z \in L$.
Prefilter F is called an n-fold positive implicative prefilter *of L, if $x^n \to (y \to z) \in F$ and $x^n \to y \in F$, then $x^n \to z \in F$, for all $x, y, z \in L$.*
Proper prefilter F is called an n-fold obstinate prefilter *of L, if $x, y \notin F$ imply $x^n \to y \in F$ and $y^n \to x \in F$, for all $x, y \in L$.*

Theorem 1. *[4] Let F be a prefilter of EQ-algebra L. Define a relation $\equiv_F$ on L as follows:*

$$x \equiv_F y \text{ if and only if } x \sim y \in F$$

Then $\equiv_F$ is an equivalence relation on L. Consider $\frac{L}{F}$ denote the quotient algebra induced by F, and $[x]_F$ denote the equivalence class of x with respect to $\equiv_F$. Moreover, if F is a filter, then $\equiv_F$ is a congruence relation on L and quotient algebra $\frac{L}{F}$ is a separated EQ-algebra.

Theorem 2. *[13] Let L be an EQ-algebra with bottom element 0. Then every n-fold obstinate proper filter of L is a maximal and n-fold positive implicative filter of L.*

Definition 4. *[5] Let L_1 and L_2 be two EQ-algebras. Then the map $\phi : L_1 \to L_2$ is called a EQ-homomorphism if and only if it satisfies the following conditions, for all $x, y \in A$:*
(i) $\phi(x \odot y) = \phi(x) \odot \phi(y)$,
(ii) $\phi(x \sim y) = \phi(x) \sim \phi(y)$,
(iii) $\phi(x \wedge y) = \phi(x) \wedge \phi(y)$.
If EQ- homomorphism ϕ is a bijection, then ϕ is called an EQ-isomorphism.

From now on, in this paper $(L, \wedge, \odot, \sim, 1)$ (or simply) L is an EQ-algebra, unless otherwise stated.

III. n-fold Integral EQ-algebras

Definition 5. *An EQ-algebra L with bottom element 0 is called an n-fold integral EQ-algebra if for all $x, y \in L$, $x^n \odot y^n = 0$, then $x^n = 0$ or $y^n = 0$.*

Example 1. *Let $L = \{0, \alpha, \beta, 1\}$ be a chain $0 < \alpha < \beta < 1$ with following tables:*

T. 1

$\odot$	0	α	β	1
0	0	0	0	0
α	0	0	α	α
β	0	α	β	β
1	0	α	β	1

T. 2

$\sim$	0	α	β	1
0	1	α	α	α
α	α	1	β	β
β	β	β	1	1
1	α	β	1	1

T. 3

$\to$	0	α	β	1
0	1	1	1	1
α	α	1	1	1
β	α	β	1	1
1	α	β	1	1

Then $(L, \wedge, \odot, \sim, 1)$ is an 2-fold integral EQ-algebra and is not 1-fold integral EQ-algebra, because $\alpha \odot \alpha = 0$ and $\alpha \neq 0$.

Example 2. *[6] Let $L = \{0, a, b, c, 1\}$ be a chain $0 < a < b < c < 1$ with following tables:*

T. 4

$\odot$	0	a	b	c	1
0	0	0	0	0	0
a	0	0	0	0	a
b	0	0	0	0	b
c	0	0	0	0	c
1	0	a	b	c	1

T. 5

$\sim$	0	a	b	c	1
0	1	0	0	0	0
a	0	1	b	b	b
b	0	b	1	c	c
c	0	b	c	1	1
1	0	b	c	1	1

T. 6

$\to$	0	a	b	c	1
0	1	1	1	1	1
a	0	1	1	1	1
b	0	b	1	1	1
c	0	b	c	1	1
1	0	b	c	1	1

Then $(L, \wedge, \odot, \sim, 1)$ is an EQ-algebra. But it is not an n-fold integral EQ-algebra, for any natural number n.

Theorem 3. *Let L be an n-fold integral EQ-algebra. Then L is $(n+1)$-fold integral EQ-algebra.*

Proof. Assume that $(x^{n+1} \odot y^{n+1}) = 0$, for $x, y \in L$. Since $(x^{n+n} \odot y^{n+n}) \leq (x^{n+1} \odot y^{n+1})$, we have $((x^2)^n \odot (y^2)^n) = (x^{n+n} \odot y^{n+n}) = 0$. Since L is an n-fold integral EQ-algebra, we get that $(x^2)^n = 0$ or $(y^2)^n = 0$. If $(x^n \odot x^n) = (x^2)^n = 0$, then since L is an n-fold integral BL-algebra, we have $x^n = 0$ and since $x^{n+1} \leq x^n$, we conclude that $x^{n+1} = 0$. Similarly, if $(y^2)^n = 0$, then $y^{n+1} = 0$. Therefore, L is an $(n+1)$-fold integral EQ-algebra.

The following example shows that the converse of Theorem 3, is not correct in general.

Example 3. *Let L be BL-algebra in Example 1. Then L is a 2-fold integral EQ-algebra and it is not a 1-fold integral EQ-algebra.*

Theorem 4. *Let F be n-fold obstinate filter of residuated EQ-algebra L. Then $\frac{L}{F}$ is an n-fold integral EQ-algebra.*

Proof. Let F be an n-fold obstinate filter of L and $([x]^n \odot [y]^n) = [0]$, for $[x], [y] \in \frac{L}{F}$. Then $[x^n \odot y^n] = [0]$ and so $(x^n \odot y^n) \sim 0 \in F$ and sin $(x^n \odot y^n) \sim 0 \leq (x^n \odot y^n) \to 0$. we have $(x^n \odot y^n) \to 0 \in F$. Since L is residuated, we get that $(x^n \to (y^n \to 0) \in F$. Now, if $x, y \in F$, then $x^n \odot y^n \in F$ and since $(x^n \odot y^n) \to 0 \in F$, we get that $0 \in F$ which it is impossible. If only $x \in F$, then $x^n \in F$, and so $y^n \to 0 \in F$ and so $y^n \sim 0 \in F$. Therefor $[y^n] = [0]$. Moreover if $y \in F$, then by similar way $[x^n] = [0]$. Now, let $x, y \notin F$, since F is a an n-fold obstinate prefilter, we have $x^n \to y \in F$ and $y^n \to x \in F$. Since by Theorem 2 F is n-fold positive implicative filter and since $x^n \to (y^n \to 0) = y^n \to (x^n \to 0) \in F$, we have

$y^n \sim 0$ and so $[x^n] = [0]$ and $[y^n] = [0]$. There for $\frac{L}{F}$ is an n-fold integral EQ-algebra.

Theorem 5. *Let A, B be two EQ-algebras, $\phi : A \longrightarrow B$ be an EQ-isomorphism and A be an n-fold integral EQ-algebra. Then $\phi(0_A) = 0_B$ and B is an n-fold integral EQ-algebra.*

Proof. The proof is straightforward.

Definition 6. *[13] Let $x \in L$. Then the* order *of x in symbol $ord(x)$ is the smallest natural number n, such that $x^n = 0$. If no such n exists, then $ord(x) = \infty$.*
EQ-algebra L with bottom element 0 is called O-local*, if $ord(x) < \infty$ or $ord(\neg x) < \infty$, for any $x \in L$ and L is called* locally finite*, if non-unite elements of L are of finite order.*

Theorem 6. *Let L be an n-fold integral spanned EQ-algebra. Then L is an O-local EQ-algebra.*

Proof. Let L be an n-fold integral spanned EQ-algebra. Then $x \odot \neg x \leq \tilde{0} = 0$, for all $x \in L$ and so $x \odot \neg x = 0$, for all $x \in L$. Hence, $x^n \odot (\neg x)^n = 0$ and since L is an n-fold integral EQ-algebra, we get that $x^n = 0$ or $(\neg x)^n = 0$. Therefore, $ord(x) = n \leq \infty$ or $ord(\neg x) = n \leq \infty$, for all $x \in L$ and so L is an O-local EQ-algebra.

Definition 7. *Let L be an EQ-algebra with bottom element 0. Then a proper prefilter F of L is called an n-fold integral prefilter, if for all $x, y \in L$,*

$$\neg(x^n \odot y^n) \in F \quad implies \quad \neg x^n \in F \quad , \quad \neg y^n \in F$$

An n-fold integral prefilter is called an n-fold integral filter, if it satisfies $(*)$.

Example 4. *[6] Let $L = \{0, a, b, 1\}$ be a chain $0 < a < b < 1$ with following tables:*

T. 7

$\odot$	0	a	b	1
0	0	0	0	0
a	0	a	a	a
b	0	a	b	b
1	0	a	b	1

T. 8

$\sim$	0	a	b	1
0	1	0	0	0
a	0	1	a	a
b	0	a	1	1
1	0	a	1	1

T. 9

$\rightarrow$	0	a	b	1
0	1	1	1	1
a	0	1	1	1
b	0	a	1	1
1	0	a	1	1

Then $(L, \wedge, \odot, \sim, 1)$ is an EQ-algebra. One can see that $F = \{1, b\}$ is an n-fold integral prefilter of L, for any natural number n.

Note that since every BL-algebra is a residuated lattice and so is EQ-algebra. Hence, we by [12], we have the following example:

Example 5. *Let $L = \{0, a, b, c, d, 1\}$ such that $0 < a < c < 1, 0 < b < d < 1, b < c$. Then L by the following operations is a BL-algebra:*

T. 10

$\odot$	0	a	b	c	d	1
0	0	0	0	0	0	0
a	0	a	0	a	0	a
b	0	0	0	0	b	b
c	0	a	0	a	b	c
d	0	0	b	b	d	d
1	0	a	b	c	d	1

T. 11

$\rightarrow$	0	a	b	c	d	1
0	1	1	1	1	1	1
a	d	1	d	1	d	1
b	c	c	1	1	1	1
c	b	c	d	1	d	1
d	a	a	c	c	1	1
1	0	a	b	c	d	1

Let $F = \{1, d\}$. Then F is a 2-fold integral filter, but F is not a 1-fold integral filter. Since $(b \odot c)^- = 0^- = 1 \in F$ and $b^- = c \notin F$, $c^- = b \notin F$.

Proposition 1. *Let L_1, L_2 be two EQ-algebras with bottom elements $0_{L_1}, 0_{L_2}$, $\phi : L_1 \to L_2$ be a EQ-homomorphism, $\phi(0_{L_1}) = 0_{L_2}$ and F be an n-fold integral prefilter of L_2. Then $\phi^{-1}(F)$ is an n-fold integral prefilter of L_1.*

Proof. The proof is straightforward.

Proposition 2. *Let L_1, L_2 be two EQ-algebras with bottom elements $0_{L_1}, 0_{L_2}$, $\phi : L_1 \to L_2$ be an EQ-isomorphism and G be an n-fold integral prefilter of L_1. Then $\phi(G)$ is an n-fold integral prefilter of L_2.*

Proof. The proof is straightforward.

Theorem 7. *Let F_1 and F_2 be two proper filter of spanned EQ-algebra L such that $F_1 \subseteq F_2$ and F_1 be n-fold integral filter of L. Then F_2 is an n-fold integral filter.*

Proof. The proof is straightforward.

Theorem 8. *Let F be an n-fold integral filter of L. Then F is an $(n+1)$-fold integral filter of L.*

Proof. The proof is straightforward.

Theorem 9. *Let F be a proper filter of L. Then F is an n-fold integral filter of L if and only if $\frac{L}{F}$ is an n-fold integral EQ-algebra.*

Proof. Let F be an n-fold integral filter of L and $[x^n] \odot [y^n] = [0]$, for $[x], [y] \in \frac{L}{F}$. Then $[x^n \odot y^n] = [0]$ and so, $\neg x^n \odot y^n) \in F$. Since F is n-fold integral filter of L. Hence $\neg x^n \in F$ or $\neg y^n \in F$. Therefore $[x^n] = [0]$ or $[y^n] = [0]$. Thus $\frac{L}{F}$ is an n-fold integral EQ-algebra.
Conversely, Let $\neg(x^n \odot y^n) \in F$ for $x, y \in L$. Then $[x^n] \odot [y^n] = [x^n \odot y^n] = [0]$ and since $\frac{L}{F}$ is an n-fold integral EQ-algebra, we conclude that $[x^n] = [0]$ or $[y^n] = [0]$, and so $\neg x^n \in F$ or $\neg y^n \in F$. Therefore, F is an n-fold integral filter of L.

Theorem 10. *Let L be spanned and separated EQ-algebra. Then the following condition are equivalent:*
(i) $\{1\}$ *is an n-fold integral.*
(ii) *Any filter of L is an n-fold integral.*
(iii) *L is an n-fold integral.*

Proof. It follows by Theorem 7 and Theorem 9.

REFERENCES

[1] P. Andrews, " An introduction to mathematical logic and type theory:to truth through proof," Kluwer, Dordrecht, 2002.

[2] M. Behzadi, L. Torkzadeh, " Obstinate and maximal prefilters in EQ-algebras," Annals of the University of Craiova, Mathematics and Computer Science Series, 44(2), pp. 228–237, 2017.

[3] R. A. Borzooei, A. Paad, " Integral filters and integral BL-algebras," Italian Journal of Pure and Applied Mathematics, 30, pp. 303–316, 2013.

[4] M. El-Zekey, " Representable good EQ-algebras," Soft. Comput., 14(9), pp. 1011–1023, 2010.

[5] M. El-Zekey, V. Novak, R. Mesiar, " On good EQ-algebras," Fuzzy Sets and Systems, 178, pp. 1–23, 2011.

[6] L. Lianzhen, X. Zhang, " Implicative and positive implicative prefilters of in EQ-algebras," Journal of intelligent and Fuzzy Systems, 26 , pp. 2087–2097, 2014.

[7] V. Nov´ak, " EQ-algebras: primary concepts and properties," Int. Joint, Czech Republic-Japan and Taiwan-Japan Symposium, Kitakyushu, Japan, pp. 219–223, 2006.

[8] V. Nov´ak, " On fuzzy type theory," Fuzzy Sets and Systems, 149, pp. 235–273, 2005.

[9] V. Nov´ak, B. De Baets, "EQ-algebras," Fuzzy Sets and Systems, 160 , pp. 2956–2978, 2009.

[10] V. Nov´ak, M. Dyba, " Non-commutative EQ-logics and their extensions in:Proc," World Congress IFSA-EUSFLAT, University of Lisbon, Portugal, 2009.

[11] A. Paad, , " Integral prefilters and integral EQ-algebras ," Mathematica Slovaca, to appear.

[12] A. Paad, R. A. Borzooei, " Generalization of integral filters in BL-algebras and n-fold integral BL-algebras," Afrika Matematika, Vol. 26, No 7, pp. 1299–1311, 2015.

[13] A. Paad, A jafari, " n-fold obstinate and n-fold fantastic prefilter of EQ-algebras," Journal of Algebra and Related Topics, Vol. 9, No 1, pp. 31–50, 2021.

[14] X. L. Xin, P. F. He, Y. W. Yang, "Characterizations of some fuzzy prefilters (filters) in EQ-algebras," The Scientific World Journal, Volume 2014, Article ID 829527, 12 pages, http://dx.doi.org/10.1155/2014/829527. 160, pp. 2956–2978, 2009.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Fuzzy Control of Autonomous Vehicle at Non-signalized Intersection in Mixed Traffic Flow

Matine Hajyan, Amirhossein Nikoofard

Department of Control and Systems, Faculty of Electrical Engineering, K.N. Toosi, University of Technology, Tehran, Iran

***Abstract*— This paper presents a control design methodology for autonomous vehicles at urban intersections based on fuzzy logic in mixed traffic flow. We have assumed there is no cooperation between an autonomous vehicle and the human-driven vehicle for passing the intersection. The autonomous vehicle is responsible for speed adaptation to prevent potential collisions using the fuzzy controller to guarantee a safe intersection crossing maneuver according to the human-driven vehicle position. We have also proposed MPC and PID controllers to simulate path tracking and evaluated both speed and turn controllers at the same time and results of these simulations show positive effects of AV in intersection control and according to RMSE results MPC controller performs effectively better than PID in path tracking.**

Index Terms — Autonomous vehicle, Non-signalized intersection, Mixed traffic flow, Fuzzy control, PID, MPC

I. INTRODUCTION

Over the past decades, the population growth and the increasing number of vehicles made traffic congestion inevitable, which is estimated 60% increase by 2030 [1]. Congestion can cause collisions, especially at significant parts of traffic networks like intersections. Unfortunately, 40% of all road injury accidents occur at intersections [2].

Many ideas and frameworks are proposed to address the need for a safer intersection. The traffic signal has always been considered a popular conventional safety method over the years. Efficiency improvement at signalized intersections is being considered widely [3, 4] but there is a significant portion of crashes that can be attributed to violating the red light; many traffic light inefficiencies can be observed during high traffic volumes [5].

With the advent of autonomous vehicles and new infrastructures, the concept of non-signalized intersection is considered [6]. Many strategies have been proposed to guarantee the safety of a non-signalized intersection in the presence of autonomous vehicles (AVs); for example, Li and Liu proposed an intersection management strategy for AV under the vehicle to infrastructure circumstance [7].

The improvement of automation and communication systems in AVs gives a new direction to intersection management research; As Belkhouche in [8] and Budan in [9] proposed collaborative methods for collision avoidance between AVs. Car following behaviors and platooning were also considered efficient approaches to achieve a reliable intersection maneuver. The main drawback of these approaches is that they require high-technology and functional materials for making infrastructures ready, which is not affordable in many countries.

Urban traffic environments will be a mixture of human-driven cars and AVs, while conventional human-driven vehicles are not going to disappear when AVs arise suddenly.

In non-cooperative approaches, AVs play the leading role of intersection control by utilizing methods like the hidden Markov model (HMM) [10], [11], or game-theory-based techniques [12]. Fuzzy logic controllers for AVs have been successfully applied to address various navigational issues [13].

Path tracking has also been considered in addition to speed control in this paper. Model predictive control (MPC), has been extensively developed especially in path tracking problems [14], [15]. Furthermore, PID as a way to achieve rapid response for attitudes is also proposed in many papers [16], [17].

The major contribution of this paper is to propose a fuzzy rule-based System (FRBS) to calculate the appropriate crossing speed at a non-signalized intersection in the presence of a human-driven vehicle to prevent likely collisions and guarantee a safe intersection maneuver. MPC and PID control methods for solving the path tracking problem of the AV are also proposed in this paper to test the speed control problem in turn situation.

The paper is organized as follows. Section I presented the problem statement and the proposed solution. In section II, the proposed fuzzy control approach are addressed. Section III discusses the structure of the fuzzy controller. Section IV introduces the MPC and PID control design for path tracking; Section V presents the simulation results, and the last section concludes the paper and proposes future works.

II. PROBLEM STATEMENT AND PROPOSED SOLUTION

We define a non-signalized intersection, which consists of two vehicles (human-driven and AV) approaching the intersection from different sides simultaneously, as depicted in Figure 1. The Av is expected to regard the priority in the situation in which a human-driven vehicle is moving toward the intersection and even stop to prevent possible crashes.
In this situation, the AV receives information about the human-driven car speed and location, and the fuzzy system analyses this information to determine the following circumstances:

- The autonomous vehicle crosses the intersection without modifying the speed.
- The autonomous vehicle gradually increases its speed in order to cross the intersection.
- The autonomous vehicle gradually decreases its speed to allow the human-driven vehicle to cross the intersection while refusing to stop.
- The autonomous vehicle should completely stop to let the other one cross.

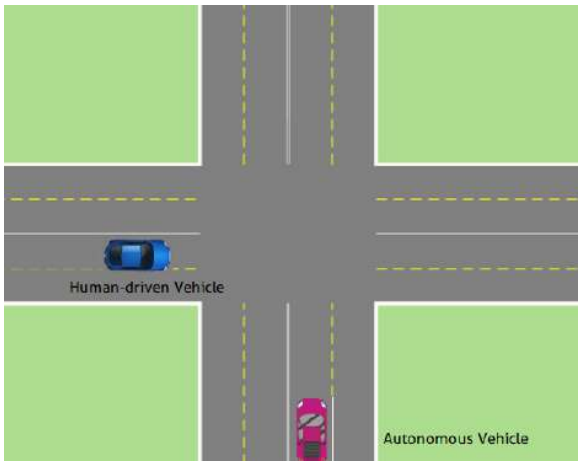


Fig. 1. Non-signalized intersection with two crossing vehicles

A further situation occurs when at least one of the vehicles is far from the intersection or has passed it; in this situation, the AV continues with normal speed. This circumstance will occur in driving in the straight line or driving in turn paths.

III. FUZZY CONTROLLER

The control system receives the position and speed of both vehicles in which (x_a, y_a) indicates the horizontal and vertical situation of the AV and (x_m, y_m) specifies the position of the human-driven one. The structure of the fuzzy controller is described as follows:

- Determines if there is an intersection situation between two vehicles or not.
- If the crossing situation was distinguished, the speed of AV would be calculated.
- Finally, the determined speed must be applied to the AV's actuators.

Each part is described in detail in the following subsections.

A. *Car position*

First, the system starts to find whether there exists the crossing situation or not, then calculates the relative position of both vehicles. Three fuzzy inputs are defined to find the exact and relative position of the cars: δ represents the angle between direction vectors of both vehicles; F_D shows the frontal deviation of the human-driven vehicle with respect to the AV; L_D shows the lateral deviation of the human-driven vehicle with respect to the AV. The variables (L_D, F_D, δ) are depicted in Fig. 2.

Therefore, the variables are calculated as follows: [10]

$$\theta_A = atan(y_A - y_A^1, x_A - x_A^1) \quad (1)$$

$$\theta_M = atan(y_M - y_M^1, x_M - x_M^1) \quad (2)$$

$$dist = \sqrt{(x_A - x_M)^2 + (y_A - y_M)^2} \quad (3)$$

$$\alpha_{rel} = 180.(atan(x_A - x_M, y_A - y_M) + \theta_A) \quad (4)$$

$$\delta = 180.(\theta_A - \theta_M)/\pi \quad (5)$$

$$F_D = dist.cos(\alpha_{rel}.\frac{\pi}{180}) \quad (6)$$

$$L_D = dist.sin(\alpha_{rel}.\frac{\pi}{180}) \quad (7)$$

$F_D > 0$ and $F_D < 0$ represent the situation in which the human-driven vehicle is moving in front or behind the AV and $F_D = 0$ shows the parallel situation of both vehicles. $L_D > 0$ and $L_D < 0$ represent the situation in which the human-driven vehicle is moving on the right or left side of the AV and $L_D = 0$ shows the situation in which the human-driven vehicle is precisely in front of or behind the AV.

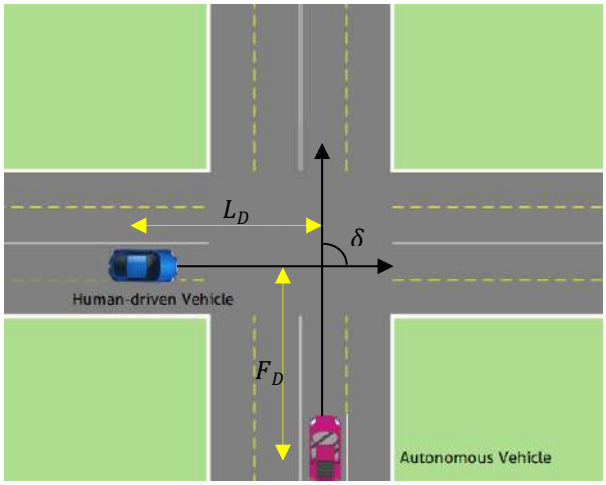


Fig. 2. Input variables

B. *Determining The Speed of The Car*

To determine the final velocity of the AV, the distance of both vehicles to the collision point is calculated due to the direction vectors of vehicles using equations (8) and (9): [10]

$$x_M + D_M \,.\, sin(\theta_M) = \; x + D_A \,.\, sin(\theta_A) \tag{8}$$

$$y_M + D_M \,.\, cos(\theta_M) = \; y_A + D_A \,.\, cos(\theta_A) \tag{9}$$

$$D_M = \tag{10}$$

$$\frac{sin(\theta_A)\,.\,(x_A - x_M) + cos(\theta_A)\,.\,(y_A - y_M)}{(cos(\theta_M)\,.\,sin(\theta_A) - cos(\theta_A)\,.\,sin(\theta_M))}$$

$$D_A = \frac{(y_A - y_M) + D_M.sin(\theta_M)}{sin(\theta_A)} \tag{11}$$

After calculating the distance to the collision point (D_M, D_A) and the velocity of each vehicle (S_M, S_A) the final velocity can be used in the fuzzy controller calculations.

Input variables (S_M, S_A, D_M, D_A) are defined by Gaussian membership functions in the fuzzy controller (each one represented by two membership functions), .and the fuzzy rule base is constructed with AND composition of all antecedents in the Sugeno fuzzy controller. Sugeno fuzzy model (TSK fuzzy model) was introduced by Takagi and Sugeno and Kang in 1685 .The general structure of the TSK model is as follows:

If "x" equals "A", AND "y" equals B, THEN
Z=f (x, y)

In this paper, the fuzzy output is the actual velocity of the autonomous vehicle.

IV. SRUCTURE OF FUZZY CONTROLLER

Fuzzy logic allows us to model different types of uncertainty and inaccuracy and it is a powerful representation tool for modeling and displaying data containing ambiguous and unreliable information of problems such as traffic issues.

The designed fuzzy controller in this work determines the final velocity of the AV using four inputs (S_M, S_A, D_M, D_A) and 24 fuzzy rules to guarantee a crossing maneuver with no collision. The diagram of the system is outlined below. In this block diagram, which is an overview of the system, the output of the fuzzy speed controller (v) and the output of PID or MPC path tracking controller (ω), are given to the car. In the path tracking controller, the difference of the real position of the vehicle (x, y, θ) and similar positions of the reference path $(x_r, y_r \theta_r)$ are calculated and represented by e_x, e_y, e_θ. Finally, it has been given to the path tracking controller to apply the appropriate steering angle (ω) to the car.

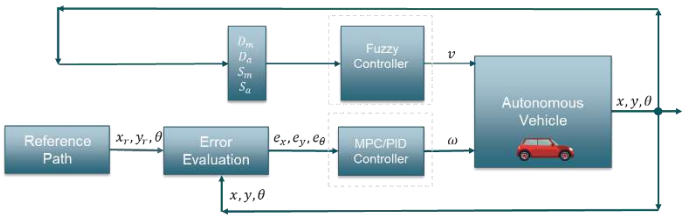


Fig. 3. System block diagram in the presence of PID or MPC controller for path tracking

Fuzzy inputs are constructed by S, Z, and Gaussian membership functions, which are well defined to make the movement smoother.

The fuzzy inputs and rules are described in the proceeding.

A. *'D_m' Input*

This fuzzy input determines the distance of the human-driven vehicle to the crossing point.

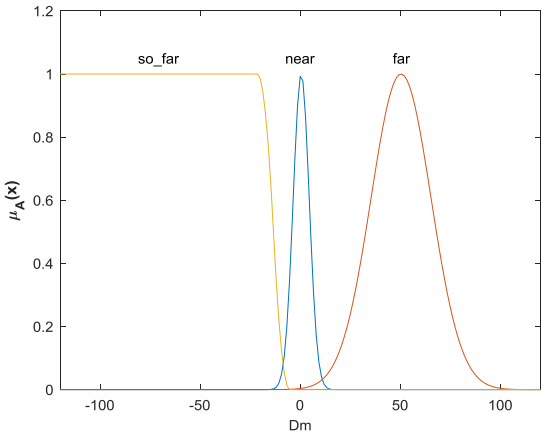


Fig. 4. Input variable "Dm"

B. *'D_A' Input*

This fuzzy input determines the distance of the AV vehicle to the crossing point**.**

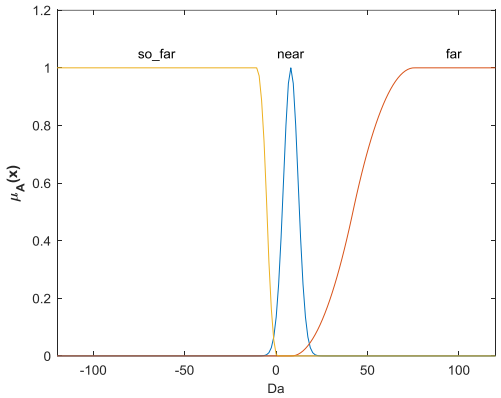


Fig. 5. Input variable “Da”

C. *'S_m' Input*

This fuzzy input determines the human-driven vehicle's speed; the speed range is defined 0-100 km/h.

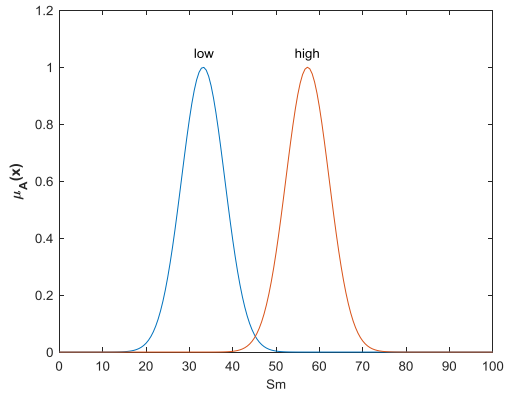


Fig. 6. Input variable “Sm”

D. *'S_A' Input*

This fuzzy input determines the AV's speed as the fourth fuzzy variable with two Gaussian membership functions; the speed range is defined as 0-50 km/h.

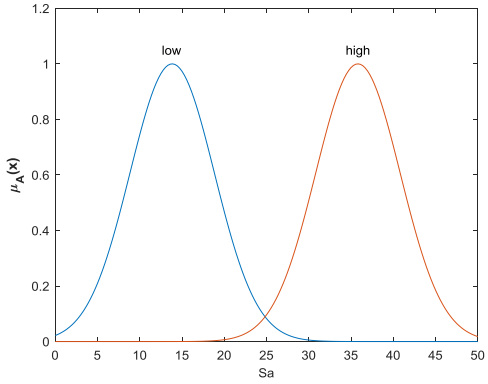


Fig. 7. Input variable “Sa”

E. *Fuzzy rules*

In this section, the if-then rule base is designed to ensure a safe and collision-free crossing situation for both vehicles; for example, one of the rules is described as follows:
If D_M is ‘near’, AND D_A is ‘near’, AND S_M is ‘high’, AND S_A is 'low', THEN V_a is ' stop'.
Three samples of proposed fuzzy rules are described in Table 1.

Table 1. The fuzzy rules to determine the AVs proper speed

Rule number	D_m	D_a	S_m	S_a	***output***
1	near	near	low	low	stop
2	near	near	high	high	stop
3	near	near	low	high	Stop
4	near	near	high	low	stop

F. *Fuzzy outputs*

Fuzzy output are defined as integer numbers in the range of 0-50 km/h; as 'stop' represents 0 km/h, 'slow' shows 20 km/h, 'middle' represents 38 km/h and 'fast' is a symbol for 50 km/h.

V. PATH TRACKING

A. *PID Controller*

In order to design a PID controller for path tracking in a turn, we calculated the difference of the real and reference positions and tuned the controller to reduce the error and propose suitable ω. The error is calculated based on the point (x_0, y_0) to line $ax + by + c = 0$ distance equation.

$$\frac{|a(x_0) + b(y_0) + c|}{\sqrt{a^2 + b^2}} \tag{12}$$

B. *MPC Controller*

Model predictive control is an appropriate control method to control the vehicles; it considers the dynamics of the vehicle and takes the path planning model and executes in each stage to consider sudden changes, such as the location of other vehicles.

Using model predictive control enables the system to control inputs based on the minimization of an objective function. This objective function is followed by constraints established by the physical constraints of the vehicle or the environment to enhance the performance of the controller.

In this research, the nonlinear predictive control algorithm (NMPC) MATLAB is used to implement a path tracking controller. The most important controller parameters are as follows in Table 2. to achieve the best performance:

Table 2. MPC controller parameters

parameter	Best Value
Ts	0.1
$prediction\ horizon$	10
$control\ horizon$	4
Q	$\begin{bmatrix} 0.2 & 0 & 0 \\ 0 & 0.2 & 0 \\ 0 & 0 & 0.2 \end{bmatrix}$
R	$\begin{bmatrix} 0.2 & 0 \\ 0 & 0.2 \end{bmatrix}$

VI. SIMULATION RESULTS

By running the simulation for different situations and modifying the speed and the distance of a human-driven vehicle to the crossing point, the proper action of the AV at interaction was observed in the face with the human-driven one.

In the crossing situation, the AV stops appropriately before the collision point and guarantees a safe crossing maneuver for both vehicles. First, the AV determines the distances to the crossing point and reduced the speed if the human-driven vehicle is close to the intersection or even stops if necessary; when the human-driven vehicle passes the intersection, the AV increases its speed and continues driving; otherwise, when the AV is not close to the crossing point, it continues with the current speed. Fig 8 - 11 illustrate the simulation results.

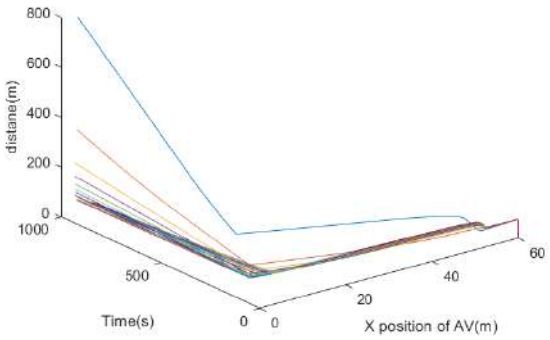


Fig. 8. Three-dimensional collision avoidance diagram of Autonomous vehicle

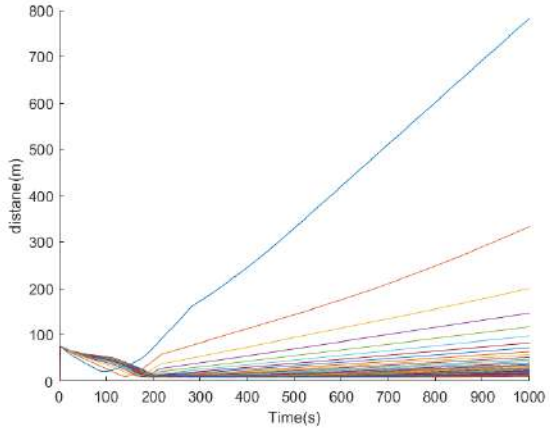


Fig. 9. Two-dimensional collision avoidance diagram of Autonomous vehicle for 50 simulations

The result of [18] in using MPC and PID in path tracking is illustrated in Fig.10 in order to compare and show the superiority of our results in path tracking.

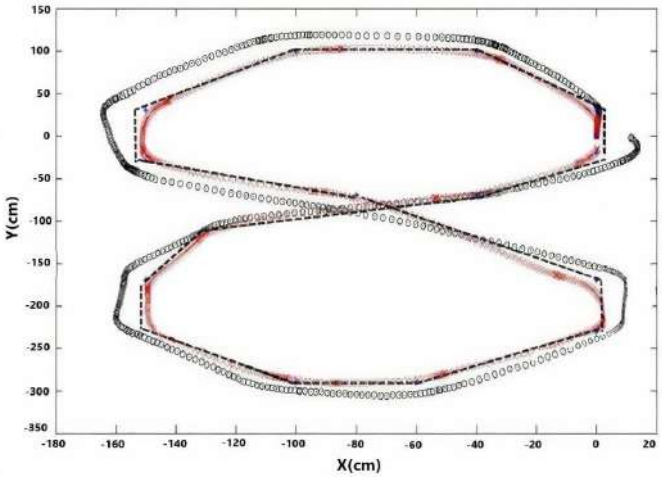


Fig. 10. Figure-of 8-shape path results for PID (empty circles) and flexible LMPC (crosses), compared to the desired path (dashes) [18]

The simulation results of PID and MPC are presented:

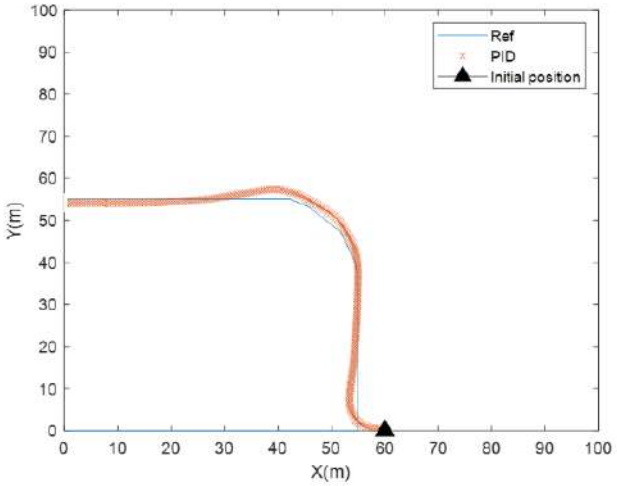


Fig. 11. Path tracking with PID controller

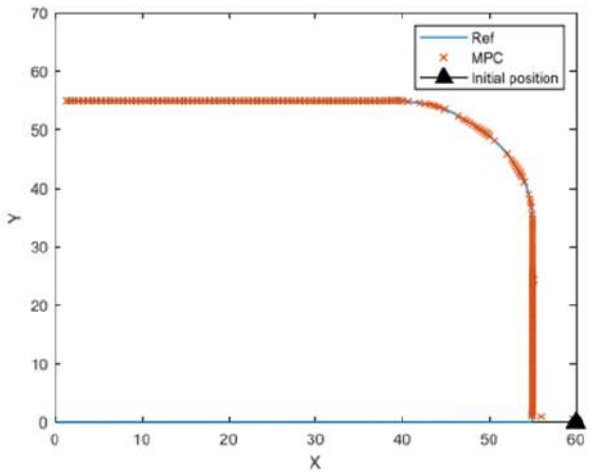


Fig 12. Path tracking with MPC controller

Unlike PID controllers, MPCs can predict future events and take control actions accordingly. In addition, MPC has the ability to deal directly with a variety of constraints on state and control inputs. In this study, the performance of both controllers in tracking is compared.

RMSE values are set for both MPC and PID controllers and for the same initial conditions the results are as follows:

MPC_{RMSE} = 1.888

PID_{RMSE} = 3.268

According to the RMS values of both controllers, the MPC controller performs better than the PID controller in path tracking.

VII. RESULTS AND FUTURE WORKS

In this paper, we proposed a fuzzy logic controller for non-signalized intersections in mixed traffic flow to secure the crossing maneuver for both human-driven and autonomous vehicles. The primary purpose of the control system is to determine the appropriate velocity of the autonomous vehicle in a situation with no cooperation between two cars; therefore, the AV is responsible for the other one's safety and should prevent collision situation in cases which the human-driver violates the intersection rules.

Moreover, the path tracking performance of the system was evaluated with both MPC and PID controllers, in which MPC assured us to have more reliable performance than PID controller.

Future works will be in the line of designing more complicated controllers for turning maneuvers at an intersection, path tracking and speed control in presence of uncertainty and implementing multi-intersection urban traffic environments.

VIII. REFERENCES

[1] Rafter, Craig, Benjamin, Anvari, Bani, Box and Simon, "Traffic responsive intersection control algorithm using GPS data," *IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, 2017.

[2] S. Noh, "Decision-Making Framework for Autonomous Driving at Road Intersections: Safeguarding Against Collision, Overly Conservative Behavior, and Violation Vehicles," *IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS*, 2018.

[3] Y. Gu, Y. Hashimoto, L.-T. Hsu, M. Iryo-Asano and S. Kamijo, "Human-like motion planning model for driving in signalized intersections," *IATSS Research*, 2016.

[4] Z. S. LUO Si-da, "Dynamic signal control for at-grade intersections under preliminary autonomous vehicle environment," *Journal of Central South University*, 2019.

[5] J. L. C. L. ELNAZ NAMAZI, "Intelligent Intersection Management Systems Considering Autonomous Vehicles: A Systematic Literature Review," *Intelligent Intersection Management Systems Considering Autonomous Vehicles*, 2019.

[6] Y. MiaomiaoLiua, "Modeling Crossing Behavior of Drivers at Unsignalized Intersections with Consideration of Risk Perception," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 45, pp. 14-26, 2017.

[7] Q. L. Yuying Li, "Intersection management for autonomous vehicles with vehicle-to-infrastructure communication," *PLoS ONE 15(7): e0235644*, 2020.

[8] F. Belkhouche, "Control of autonomous vehicles at an unsignalized intersection," *American Control Conference*, 2017.

[9] G. Budan, K. Hayatleh, D. Morrey, P. Ball and F. Shadbolt, "An analysis of vehicle-to-infrastructure communications for non-signalised intersection control under mixed driving behaviour," *Analog Integrated Circuits and Signal Processing*, 2018.

[10] E. Onieva, V. Milanés, J. Villagrá , J. Pérez and J. Godoy, "Genetic optimization of a vehicle fuzzy decision system for intersections," *Expert Systems with Applications* , 2012.

[11] G. X. a. H. C. Weilong Song, G. Xiong and H. Chen, "Intention-Aware Autonomous Driving Decision-Making in an Uncontrolled Intersection," *Mathematical Problems in Engineering*, 2016.

[12] M. Elhenway, A. A. Elbery and A. A. Hassan, "An Intersection Game-Theory-Based Traffic Control Algorithm in a Connected Vehicle Environment," *IEEE 18th International Conference on Intelligent Transportation Systems*, 2015.

[13] V. Stoian, "A Control Algorithm for Autonomous Electric Vehicles by Fuzzy Logic," *Advanced Engineering Forum 27:103-110*, 2018.

[14] Y. G. M. Q. Shuyou Yu, "MPC for path following problems of Wheeled Mobile Robots," in *IFAC (International Federation of Automatic Control)* , 2018.

[15] Y. X. S. M. I. L. D. a. P. C. Zhongqi Sun, "Tracking of unicycle robots using event-based MPC with adaptive prediction horizon," *IEEE/ASME Transactions on Mechatronics*, vol. 25, no. 2, pp. 739 - 749, 2019.

[16] P. Damodaran, "Design and Implementation of GA Tuned PID Controller for Desired Interaction and Trajectory Tracking of Wheeled Mobile Robot," in *the Advances in Robotics*, 2017.

[17] B. Zhao, H. Wang, Q. Li, J. Li and Y. Zhao, "PID Trajectory Tracking Control of Autonomous Ground Vehicle Based on Genetic Algorithm," *Chinese Control And Decision Conference (CCDC)*, 2019.

[18] N. L. Lluis Pacheco, "Testing PID and MPC Performance for Mobile Robot Local Path-Following," *ARTICLEInternational Journal of Advanced Robotic Systems*, vol. 12 , no. 11, 2015.

[19] D. Qun Cao, "Self-triggered MPC for trajectory tracking of unicycle-type robots with external disturbance," *Journal of the Franklin Institute* , vol. 356, no. 11, pp. 5593-5610, 2019.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Fuzzy Topic Modeling On Persian News

1st Vahid Heidari, and 2nd Seyed Mahmoud Taheri

1st School of Engineering Science, Univercity of Tehran, Tehran, Iran, vahid.heidari@ut.ac.ir
2nd School of Engineering Science, Univercity of Tehran, Tehran, Iran, sm_taheri@ut.ac.ir

***Abstract*— In this paper, we investigate two versions of unsupervised clustering Latent Dirichlet Allocation (LDA) methods in original and fuzzy forms. Gibbs samplers are implemented for examining clustering performances on Persian news dataset. Our experimental results are showed that the fuzzy implementation of LDA performs better in text clustering tasks.**

***Index Terms*— Fuzzy Bag of Words, Fuzzy Latent Dirichlet Allocation (FLDA), Natural Language Processing, Text Mining.**

I. Introduction

The way we access to news is changing every day. Specially, the emergence of portable cell phones made the news more accessible than ever. Recent study of ISPA [1] showed that popularity of online sources are growing in comparison to offline sources in period of three years leading to 2021 in Iran. These online resources provide rich textual datasets for researchers in natural language processing (NLP).

The incredible size of text contents, which are created by users every day on the internet, implies unsupervised methods of machine learning for clustering these documents. One approach is to arrange documents with common subjects in groups or clusters. These hidden subjects are called "topics" and this method is called topic modeling [1].

One of the most powerful topic modeling methods is Latent Dirichlet Allocation (LDA) [2]. LDA has been successfully applied to cluster documents. However, it suffers from three major problems [3]:

1) Wrong importance assignment of words,
2) High sparsity of documents,
3) Ignoring semantic information.

These issues are caused by using Bag-of-Word (BoW) for document model. In classical BoW model, each word is a unit vector with only **one** element is equal to 1 and others are 0, this representation is known as one-hot vector [4].

The fuzzy Bag-of-Word (FBoW) addressed these issues. It replaced the one-hot vectors with more dense and robust document representation, via word embedding. Word embedding is a process of assigning a dense and low dimensional vector to each word, such that words with similar semantic meaning are placed closed together in vector space. For further reduction of dimensionality and sparsity, in a large corpus, only high-frequency words are usually selected as basis terms, therefore $T = \{t_1, \cdots, t_B\}$ is a subset of vocabulary, i.e. $T \subset V$ [4].

Fuzzy representation of documents in FBoW, leads to fuzzy LDA (FLDA) implementation. Because FBoW resolves three problems of traditional BoW, then FLDA performs better than LDA. The first and second problems have been solved by assigning fuzzy weight of all basis terms to all words in a document. Also, the third problem has been solved, since fuzzy representation considers semantic association among words automatically [3].

In this paper, we compare original LDA to FLDA based on accuracy of clustering, and show that the FLDA has better results than LDA. In Section II, we review a few topic modeling methods and describe Gibbs samplers for LDA and FLDA. Section III presents our dataset preparation procedure (collected on 24th November 2021) and evaluation measure. Section IV describes our experiment settings and results. Section V includes conclusion and some directions for future works.

II. Related Works

In text mining, the "topic" is a short description for a large collection of discrete data such as text document. The goal of topic modeling is to find a probabilistic model of text corpora, that enables us to process these large datasets in a more efficient way [5].

Most important applications of topic modeling are information retrieval and knowledge discovery. Fuzzy set theory approach to topic modeling is a novel method for discovering hidden knowledge in extensively large datasets. For example, in [6] authors proposed a fuzzy

[1] http://www.ispa.ir/Default/Details/fa/2338/

clustering method, called Fuzzy Latent Semantic Analysis (FLSA), based on Fuzzy C-Means (FCM) method. FLSA tries to avoid redundancy problem, which most medical text documents have redundant notes. The redundancy has negative impact on topic modeling.

LDA is a popular technique of topic modeling. LDA is a generative model for creating a text corpus. It is well known that the posterior of this model is intractable, so we need an approximation scheme to estimate its parameters. Many approximation schemes have been proposed such as Gibbs sampling through Markov Chain Monte Carlo (MCMC) algorithm, Expectation Maximization (EM), Variational Bayes (VB) inference. A survey of methods for approximating the posterior of LDA, as well as applications of LDA in other fields of science, can be found in [2].

Original LDA uses Variational Bayes (VB) method via EM algorithm to estimate model parameters [5]. The main idea behind VB method is to replace the intractable posterior distribution by tractable family of distributions. The Mean-Field approximation is adopted to design factorizable variational distributions, and variational parameters are chosen by minimizing Kullback-Leibler (KL) divergence between variational distributions and the true posterior [5]. The Gibbs sampling is another way of approximating intractable posteriors, when sampling from posterior is not applicable but sampling from full conditional distributions of a latent variable given all other variables can be computed. In the case of LDA, a Gibbs sampler can be designed in a more efficient way to reach a faster convergence rate. Because of conjugacy of prior distributions, parameters of joint distribution can be integrated out. A designer can use this fact to build an efficient algorithm by sample from distribution of latent variables directly. This type of sampler is called Collapsed Gibbs sampler [7]. In this paper, we examine Gibbs sampling implementations of LDA and FLDA, because it is straight forward and easy to implement.

A. Gibbs Sampler for LDA

The goal of LDA is to capture distribution of words in each topic, $\Pr(w_{dn} \mid z_{dn}, \phi)$, and distribution of topics per document, $\Pr(\theta_d \mid \alpha)$ in a corpus C containing D documents. A Gibbs sampler for LDA, includes three steps as follows

$$\Pr(z_{di} = k \mid w_{di}, \phi_{kv}) \propto \frac{n_{dk} + \alpha}{\sum_{k'=1}^{K} n_{dk'} + \alpha} \times \frac{m_{wk} + \beta}{\sum_{w'=1}^{V} m_{w'k} + \beta} \tag{1}$$

$$\theta_{dk} = \frac{n_{dk} + \alpha}{\sum_{k'=1}^{K} n_{dk'} + \alpha} \tag{2}$$

$$\phi_{kw} = \frac{m_{wk} + \beta}{\sum_{w'=1}^{V} m_{w'k} + \beta}, \tag{3}$$

where, K is number of topics, V number of word in vocabulary, n_{dk} is the number of words in a document d which are assigned to k^{th} topic, m_{wk} is the number of assignments of a word w to k^{th} topic, and α and β are hyper-parameters of prior distributions. The Gibbs sampler, draws samples from equations (1), (2), and (3) in a loop until convergence or reaching to maximum running time limitation [8], [9]. The graphical model of LDA is depicted in Figure 1.

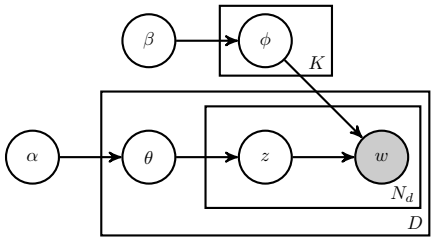


Fig. 1: LDA Graphical Model [7].

The probability distribution of words in k^{th} topic is specified by the k^{th} row of ϕ matrix, i.e. ϕ_k. The ϕ matrix is global to the corpus, so that ϕ is shared among all documents. The θ_d is a K-elements vector of topic distribution of d^{th} document. The z_{di} is an indicator latent variable, indicating topic of i^{th} word in document d. The θ_d is local to document d, because it is sampled once for each document, and the z_{di} is local to each word [7]. The generative process of creating a corpus of documents in LDA is summarized in Algorithm 1.

Algorithm 1 Generative Process for Corpus

Select word distribution matrix $\phi_k \sim \mathcal{D}irichlet(\beta)$.
for $d = 1, \cdots, D$ **do**
  Draw sample for $\theta_d \sim \mathcal{D}irichlet(\alpha)$.
  for $i = 1, \cdots, N_d$ **do**
    Draw sample for $z_{di} \sim \mathcal{C}ategoritcal(\theta_d)$.
    Draw sample for $w_{di} \sim \mathcal{C}ategoritcal(\phi_{z_{di}})$.

B. Gibbs Sampler for FLDA

FLDA replaces one-hot vector representation of words by a lower dimensional and dense vector, and a document is represented as $f = [f_1, \cdots, f_B]$, where B is the number of basis terms. The f_i value is obtained as follows

$$f_i = \sum_{j=1}^{N_d} \mu_{t_i}(w_j) \cdot c_j,$$

where N_d is the number of words in document d, and c_j is the frequency of word w_j in that document [3]. The membership function, μ_{t_i}, is cosine similarity between words, defined as:

$$\mu_{t_i}(w_j) = \begin{cases} \cos(W[t_i], W[w_j]) & if \cos(W[t_i], W[w_j]) > 0 \\ 0 & otherwise \end{cases},$$

where W is word embedding matrix of our vocabulary, and $t_i \in T$ is a basis term. Only positive cosine similarity is considered. When two words are similar in meaning, the

corresponding vectors in embedding space are close, then the cosine similarity is close to 1, and cosine similarity is negative when two words have semantically differece meaning. There are several methods to find word embedding matrix. In this paper we use "skip-gram" language model of "word2vec" to find dense representation of all words [10].

For example, the word "Computer" is similar to words "Desktop" and "Mouse', but dissimilar to words "Government" and "International" in word embedding space. A one-hot representation of these words is as follows

$$\begin{aligned}\text{Computer} &= (1,0,0,0,0,0,0)\\ \text{Desktop} &= (0,1,0,0,0,0,0)\\ \text{Mouse} &= (0,0,1,0,0,0,0)\\ \text{Panasonic} &= (0,0,0,1,0,0,0)\\ \text{Government} &= (0,0,0,0,1,0,0)\\ \text{International} &= (0,0,0,0,0,1,0)\\ \text{Found} &= (0,0,0,0,0,0,1).\end{aligned}$$

In this representation the cosine similarity between each pair of words is always 0. After performing word embedding, a 2 dimensional representation might be as Figure 2. And we can compute cosine similarity of words. As an example, the cosine similarity of word "Computer" and other words is showed in Table I.

After obtaining word embedding of vocabulary we can represent Gibbs sampler for FLDA. Steps of Gibbs sampler for FLDA is same as LDA, except counting values of n_{dk} and m_{wk} are fuzzy in the following way:

$$\Pr(z_{di}=k \mid w_{di},\phi_{kv}) \propto \frac{\sum_{v=1}^{n_{dk}} f_{dv}+\alpha}{\sum_{k'=1}^{K}\sum_{v=1}^{n_{dk'}} f_{dv}+\alpha} \times \frac{\sum_{v=1}^{m_{wk}} f_{d_v v}+\beta}{\sum_{w'=1}^{V}\sum_{v'=1}^{m_{w'k}} f_{d_{v'}v'}+\beta} \tag{4}$$

$$\theta_{dk} = \frac{\sum_{v=1}^{n_{dk}} f_{dv}+\alpha}{\sum_{k'=1}^{K}\sum_{v=1}^{n_{dk'}} f_{dv}+\alpha} \tag{5}$$

$$\phi_{kw} = \frac{\sum_{v=1}^{m_{wk}} f_{d_v v}+\beta}{\sum_{w'=1}^{V}\sum_{v'=1}^{m_{w'k}} f_{d_{v'}v'}+\beta}, \tag{6}$$

where f_{dv} is membership degree of basis term and word v, d_v is a document containing word v, ϕ_{kw} is membership degree that word w belongs to topic k, and θ_{dk} is membership degree of document d belongs to topic k [4]. In fuzzy setting, the Gibbs sampler algorithm is summarized in Algorithm 2.

TABLE I: Cosine Similarity of word "کامپیوتر" (Computer) with some other words

Word	Meaning	Cosine Similarity
دسکتاپ	Desktop	0.997
ماوس	Mouse	0.882
پاناسونیک	Panasonic	0.833
صندوق	Fund	−0.018
بین‌المللی	International	−0.034
دولت	Government	−0.035

Algorithm 2 Gibbs Sampler Algorithm

Initialize Z randomly.
Fill count matrices based on Z assignments.
repeat
  for d=1 to D **do**
    for i=1 to B **do**
      Draw topic assignment z_{di} from (4).
      Update fuzzy count matrices.
  Draw samples for θ and ϕ from (5), and (6) respectively.
until Convergence or reaching to maximum running time limitation

III. Materials and Methods

In this section, we explain data preparation and evaluation measures. Before crawling, we selected only two **Economics** and **Politics** news categories, because these articles are analytical and have longer text bodies in comparison to other categories.

A. Data Preparation

To examine the performance of clustering documents, we collected dataset from online sources. These raw inputs need to be processed into appropriate forms for running various text clustering tasks. Our experiment includes the following steps:

1) Collecting dataset (list of news, news pages) by crawling YJC [2] website,
2) Extracting content of news body from HTML pages,
3) Removing stop words,
4) Training skip-gram model of word2vec to generate word embedding matrix,
5) Clustering news pages using LDA and FLDA methods,

[2] https://yjc.ir

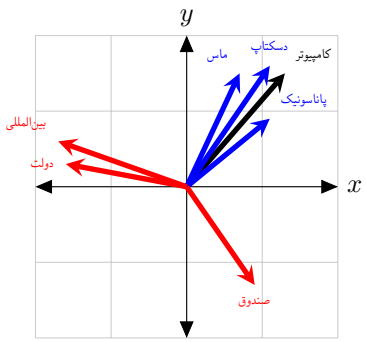


Fig. 2: Word Vectors in Embedding Space

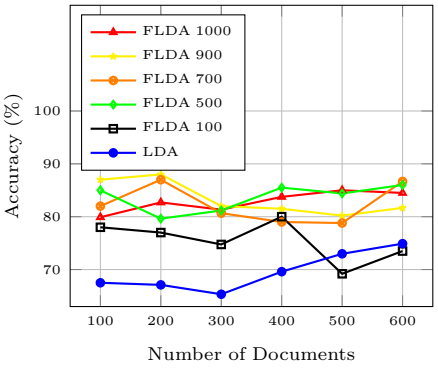


Fig. 3: Accuracy of Document Clustering of LDA and FLDA with Different Number of Basis Terms (B)

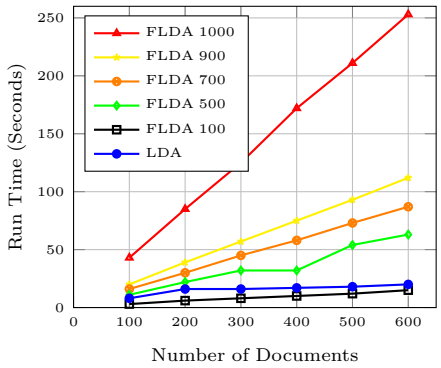


Fig. 4: Run Time Comparison between LDA and FLDAs with Different Number of Basis Terms

6) Comparing accuracy of clustering.

After downloading pages and extracting news body contents, we have a corpus with $D = 1374$ documents, and $V = 14315$ unique words.

B. Evaluation Measure

There are some measures such as Purity, Rand Index, Mutual Information, and F measure for comparing clustering performance. In this paper, we have labeled data, so we use an Ad-Hoc method to measure clustering accuracy. After clustering, we can assign tags to each clusters in arbitrary order. There are $n = 1, \cdots, K!$ permutations of tag assignments $P_n = (p_1, \cdots, p_K)$, $p_i \in \{1, \cdots, K\}$. Counting number of tag assignments equal to true labels, we can compute accuracy of n^{th} permutation as:

$$\text{Accuracy}_n = \frac{\sum_{d=1}^{D} \mathbb{I}(\arg\max_k \theta_d = C_n(d))}{D} \times 100, \quad (7)$$

where $\mathbb{I}(\bullet)$ is indicator function, and $C_n : d \to p$ is the tag assignment of cluster, which d^{th} document is placed into. This method is computationally feasible, until the number of clusters is kept small (unless, other measures can be used). The best tag assignment is a permutation with the highest accuracy.

IV. Experiments and Results

In this section, we designe experiments to examine the performance of LDA and FLDA models. First, we investigate clustering capability of these models based on accuracy. Second, we analyze time complexity and execution time in different configurations.

A. Accuracy of Clustering

We conduct text clustering experiments on news corpus collected by the procedure mentioned in Section III-A. We gradually increase number of documents in corpus, and measure accuracy of clustering by calculating relation (7). We report average accuracy after running each test 10 times. The values of hyper-parameters $\alpha = 0.5$ and $\beta = 0.01$, word embedding matrix dimension is 14315×100, and number of clusters $K = 2$. In each test we select basis terms of size $B \in \{100, 500, 700, 900, 1000\}$, then we give them as an input to FLDA. We repeat tests with different number of documents. Figure 3 presents the performance of document clustering in this particular dataset and configurations.

B. Time Complexity

In conventional LDA each iteration needs $N_d K + K$ operations for sampling the latent variables z_d and θ_d related to d^{th} document. Sampling would be need to be done for all documents in the corpus, so the whole process is multiplied by the factor of M. It is a reasonable assumption to replace N_d by average number of words in documents, say N. Sampling of the ϕ parameter needs $MKVN$ operations to be done. In practice, the average execution time of this algorithm is in order of $O(MKN^2)$ to complete one iteration of sampling from all of the latent variables and parameters. We can conclude that run time complexity of LDA linearly increases by the number of documents (M).

In the case of FLDA, we have another additional parameter B, which is defined as the number of basis terms, which could have significant impact on running time of this algorithm. By increasing the number of basis terms, the running time increases.

To investigate the effect of parameter B on execution time of FLDA, we designed the following experiment. We chose basis terms of size $B \in \{100, 500, 700, 900, 1000\}$, then we gradually incremented the number of documents in corpus. Then we depicted the execution time of algorithm in Figure 4 in comparison to LDA.

V. Conclusion

Fuzzy set theory models membership of documents and words by changing perspective of reasoning from two numbers 0 and 1 to values in unit interval $[0, 1]$. Therefore, the vagueness of natural language is more precisely modeled in topic modeling application.

Transforming one-hot word vectors to dense vectors, and using fuzzy representation of documents, improves

performance of LDA in document clustering application. According to Figure 3, the FLDA accuracy is 10% higher than LDA overall.

The advantage of using basis terms is that FLDA can perform better than LDA. Because basis terms provide more information to FLDA which is inaccessible to LDA. Therefore, FLDA can be categorized as semi-supervised method. FLDA performance can be further improved by using basis terms that offered by domain experts of each field, instead of simply choosing top frequent words of previously clustered documents by LDA or other methods.

Figure 4 shows that both running time of FLDA and LDA are linear in terms of number of documents, but FLDA running time slop is higher than LDA, because of larger constant in big 'O' notation time complexity, caused by the additional B parameter.

As mentioned in Section II-B, LDA is a generative model. We found that, this property is very appealing, because we can make simulated datasets in various parameter setups and sizes. During implementation of MCMC algorithm, testing and troubleshooting of smaller dataset is more convenient, before trying real datasets.

Finally, possible directions for future works are as follows:

1) Examining performance of proposed method on bigger datasets,
2) Further clustering improvement by using user comments about news,
3) Automatic extraction of basis terms during execution of FLDA, instead giving them as predefined ones,
4) Deeper investigation into the effect of additional parameters in FLDA on accuracy of clustering.

It should be noted that source codes are accessible at TopicModelingFuzzy [3] .

References

[1] J. Rashid, S. M. A. Shah, and A. Irtaza, "Fuzzy topic modeling approach for text mining over short text," *Information Processing & Management*, vol. 56, no. 6, p. 102060, 2019.

[2] H. Jelodar, Y. Wang, C. Yuan, X. Feng, X. Jiang, Y. Li, and L. Zhao, "Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey," 2018.

[3] N. Akhtar, M. Beg, and H. Javed, *Topic Modelling with Fuzzy Document Representation*. Springer Singapore, 2019, pp. 577–587.

[4] R. Zhao and K. Mao, "Fuzzy bag-of-words model for document representation," *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 2, pp. 794–804, 2018.

[5] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.

[6] A. Karami, A. Gangopadhyay, B. Zhou, and H. Kharrazi, "Fuzzy approach topic discovery in health and medical corpora," *International Journal of Fuzzy Systems*, vol. 20, no. 4, pp. 1334–1345, 2017.

[7] C. Geigle, "Inference methods for latent Dirichlet allocation," 2016.

[8] T. Griffiths, "Gibbs sampling in the generative model of latent Dirichlet allocation," 2002.

[9] T. L. Griffiths and M. Steyvers, "Finding scientific topics," *Proceedings of the National Academy of Sciences*, vol. 101, no. suppl 1, pp. 5228–5235, 2004.

[10] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," 2013.

[3] https://github.com/VahidHeidari/TopicModelingFuzzy

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

How weight-sharing mechanisms affect the performance of deep Siamese networks

Mahvash Mohazzebi, Mahdi Eftekhari, Mahdi Shariatzadeh

Department of Computer Engineering, *Shahid Bahonar University of Kerman*, Kerman, Iran
mohazzebi@eng.uk.ac.ir, m.eftekhari@uk.ac.ir, shariat@eng.uk.ac.ir

***Abstract*—Siamese networks usually consist of two parallel networks and each network is made of a backbone and some fully connected layers. The entire Siamese structure is trained by a loss function, while the similarity of the latent spaces of the two networks is decisive at the inference time. The weights of the network can be shared in Siamese networks. Therefore, in this paper, four weight-sharing mechanisms are proposed, and their effects on the performance of Siamese networks are investigated. These methods are named "full weight sharing," "partial weight sharing in FC," "partial weight sharing in CNN," and "without weight sharing." In this paper, four network backbones are examined and trained by employing the four aforementioned weight-sharing methods; two losses (proxy anchor loss and contrastive loss) are utilized for training. According to the obtained results, "partial weight sharing in FC" combined with contrastive loss has the best result on all backbones.**

***Index Terms*—Siamese network, weight sharing, proxy anchor loss, contrastive loss**

I. Introduction

Siamese networks, which are a special type of convolutional neural network, are an emerging technique based on deep learning. Neural networks that consist of two or more similar architectures and have weight sharing between them are called Siamese neural networks [1]. These networks are used to learn the similarity between the input data [2]. These networks map the input data to an embedding space and try to train the network using various loss functions [3].

In general, the goal of Siamese neural networks is not to classify but to learn how to encode images and then determine the differences between the input images [4]. In this type of network, the difference between images is calculated by minimizing the distance between similar inputs and maximizing the difference between dissimilar inputs.

In the next section, the related works are reviewed. Section III presents the experimental setup in four subsections. Section IV deals with the approach. Section V presents the experimental results. In section VI, the conclusion is made.

II. Related works

One of the applications that can be implemented using Siamese networks is identifying kinship relationships [5]. Due to the age differences and gender differences between parents and children, in identifying kinship relationships, Siamese networks are used to identify similarities between children and parents. Yu et al. [5] made two suggestions. In the first suggestion, the image of one parent and the image of the child are given as input to the Siamese network, and the similarity between these images is calculated. According to the second suggestion, the relationship between the two parents and the child, in which each parent and child is given to a network, and the degree of similarity between child and mother and the degree of similarity between child and father are calculated separately; finally, the similarity between the two is decided.

Humans pay more attention to the differences between two faces to distinguish them. Inspired by how humans recognize two different faces, Han et al. [6] proposed a new structure for convolutional neural network networks and face recognition tasks. This structure receives two different images of the face as input. First, the main features of the face are extracted for each input, and then based on these features, the kernel generator, which consists of several sub generators, obtains a personal kernel for each input image, and based on these two personal kernels, the similarity of two input images is calculated.

Dong and Shen [7] tried to use the best triple inputs to find effective and powerful features by introducing a new triplet loss function. Their goal was to increase the joint probability between all three inputs during training. They were able to use this technique for the object tracking problem.

Language processing problems can usually be grouped using semantic matching. These problems are limited to short sentences or a paragraph due to problems such as the quadratic computational complexity of the models that work on these problems. Yang et al. [8], offering several innovations and using the structure of Siamese networks, provided an encoder for long-form document matching. This model can increase the maximum length of input words from 512 to 2048.

Due to the architecture and features of the Siamese net-

works, these networks are instrumental in locating objects. Zheng and Peng [9] found that the depth of the networks used in the Siamese backbone is not effective in achieving better accuracy. To improve accuracy, they introduced two new structures that, while lightweight, perform very fast real-time tracking operations.

Different losses are used in Siamese networks. One of the most famous is contrastive loss, which has problems. Cheng et al. [10] proposed a modified contrastive loss called MCL to overcome contrastive loss problems such as overfitting and low sampling. By using this loss, in addition to increasing accuracy, they have also made significant improvements in computational costs. In addition to loss, they introduced a new method. In their method, several ResNets are first trained based on softmax or MCL losses and then combined. Finally, using the joint Bayesian method, they integrate more features for more learning.

He et al. [11] used Siamese networks to identify faces. They used binary cross-entropy as the loss function, and the results obtained from this network structure were much better than the case where the standard sigmoid function is the loss function. The general structure is that each branch of the Siamese network is given an image, and these images are passed through the Convolutional Neural Network (CNN); the output of each branch is a vector. By using the cross-entropy loss function, the network is trained with the goal of setting the network output to 0 for similar images and 1 for dissimilar images.

Signature is a user authentication tool. So far, many methods have been proposed to verify signatures. Gosh et al. [12] used a Siamese network to recognize 3D signatures. In this structure, a signature is selected as an anchor, and two other signatures, one from the same class as the anchor and the other from a different class, are given to the network, and finally, the similarity is calculated using the triplet loss function. Chakladar et al. [13] proposed a method for verifying signatures using Siamese networks. The input of each branch of the network includes a signature and its Electroencephalogram (EEG) signal. Each EEG signal passes through the LSTM network, and the signed image passes through the CNN network. Finally, the compatibility function between the outputs of these two networks is calculated. Weights are shared between the two branches of the Siamese network. The final output of the Siamese network includes the similarity between the compatibility function of the two branches of the Siamese network. In this structure, the contrastive loss function is used to calculate similarity. With this new structure, 98.57% accuracy has been achieved.

Siamese networks suffer from the possibility of collapse [14]. Chen et al. [15] were able to prevent the collapse of Siamese networks by using two different augmentations of an image given to the Siamese network and using a stop-grad system. Chen et al. [16] also used a Siamese network. Their network structure was such that it repelled different images from different classes and, at the same time, absorbed images with the same class. In their proposed architecture, the batch size for training should be large to prevent collapsing.

III. Experimental setup

A. Evaluation measures

In this study, performance measures for evaluating Siamese networks are introduced. Three criteria are used to evaluate these networks, including accuracy, precision, and recall, which are defined as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

where TP stands for true-positive, TN stands for true-negative, FP stands for false-positive, and FN stands for false-negative in the confusion matrix.

B. Data sets and parameters

In this study, Siamese networks were trained to perform experiments using the AT&T face database and the fingerprint dataset. The AT&T face database contains 400 grayscale images of 40 different people, with ten images per person in different situations. In the experiments, five images of each person, a total of 200 images, were considered as training data. Similarly, three images of each person were provided for validation data, and two images of each person were provided for test data. Validation data contained 120 images, and test data contained 80 images. In addition to the AT&T face database, the fingerprint dataset was also used. This dataset contains 800 grayscale fingerprint images belonging to 10 different classes. Each class contains 80 images of a fingerprint obtained from different angles and with different resolutions. Fifty fingerprints from each class were selected as training data. There were a total of 500 fingerprints in the training set. The number of fingerprints for validation data was 200. Twenty fingerprints for each class, and, the number of fingerprints for test data was 100, ten fingerprints for each class. Networks were trained in 200 epochs. It is important to note that trained models were stored and used for the most accurate state of validation data. In all experiments, a batch size $b = 64$ was used. The AdamW optimizer was used to train the models. Learning rate $lr = 0.0001$ and embedding size $e = 8$ was considered.

C. Loss functions

The experiments performed in this study are on a Siamese network using contrastive loss and proxy anchor loss. Loss function can be categorized into two categories: pairwise-based and proxy-based. Pairwise-based losses consider the relationships between both data and are constructed based on the distance between the data. The disadvantages of these losses include high training complexity, low convergence speed, and high probability of over-fitting. Proxy-based losses, because they use the concept of the proxy, do not have the problem of training complexity and converge faster. Proxy means a

representative of a part of training data. In this type of loss, data is compared to proxies instead of data. The disadvantages of these losses include ignoring the relationship between data and data.

Proxy anchor loss is in the category of proxy-based losses but uses the positive features of the two categories discussed above. Because it uses a proxy, this type of loss has a high convergence speed and is also less complex. This loss maintains the data relationship with the proxy because it compares proxies with data. The proxy anchor routine is to first assign a proxy to each class and pull data from the same class to the proxy. The farther the data is from the proxy, the more powerful it is. Data that is different from the class repels the proxy. The closer the data is to the proxy, the more power it will have to repel the proxy [17].

Proxy anchor loss is calculated by the following equation:

$$L(X) = \frac{1}{|P^+|} \sum_{p \in P^+} log(1 + \sum_{x \in X_p{}^+} e^{-\alpha(s(x,p)-\delta)}) + \frac{1}{|P|} \sum_{p \in P} log(1 + \sum_{x \in X_p{}^-} e^{-\alpha(s(x,p)+\delta)}), \tag{4}$$

where s denotes cosine similarity. δ represents the margin. α is the scaling factor. δ and α values must be positive. P stands for a set of proxies. P^+ stands for positive proxies and p stands for any proxy. X represents the embedding space of the data in the batch. ${X_p}^+$ stands for positive embedding vectors of p and ${X_p}^-$ is defined as ${X_p}^- = X - {X_p}^+$.

An example of a pairwise loss is contrastive loss. Contrastive loss as the following equation tries to create a distance of m between dissimilar data [18]. In equation 5, similar data with $Y = 0$ get closer to each other.

$$L(Y, X_1, X_2) = (1 - Y)\frac{1}{2}(D)^2 + (Y)\frac{1}{2}\{max(0, m - D)\}^2, \tag{5}$$

where L is the contrastive loss, X_1 and X_2 are data, and Y is the label of the two data, so that if the two data belong to the same class, $Y = 0$, otherwise $Y = 1$. m is the margin and is $m > 0$. D is the Euclidean distance.

D. Backbones

SqueezeNet is a CNN with complicated architecture. The network was introduced in 2016 by researchers at California, Berkeley, and Stanford Universities [19].]. The goal was to introduce a network with fewer parameters to take up less memory. SqueezeNet can reach the accuracy of the AlexNet network in the ImageNet database if the AlexNet parameters are 50 times higher but almost three times slower. The SqueezeNet network consists of "squeeze" and "expand" layers. So far, many models have been developed, such as SqueezeDet for object detection on images, SqueezeSeg for semantic segmentation, SqueezeNext for image classification, and SqueezeNAS for Fast Neural Architecture Search for Faster Semantic Segmentation. In this article, squeeznet is used as a backbone in the branches of the Siamese network.

GoogleNet was introduced in 2014 by Google researchers [20]. GoogleNet wins the ILSVRC 2014 image classification challenge. This network is a CNN with a complex architecture based on the architecture and structure of inception. This network is suitable for detection, recognition, and image classification tasks. GoogleNet can create a more profound architecture due to the use of different methods such as convolution 1×1 and global average pooling. GoogleNet has 12 times fewer parameters than AlexNet, so it requires less memory. GoogleNet has 22 deep layers with 27 pooling layers. In this article, GoogleNet is used as a backbone in the branches of the Siamese network.

ResNet is one of the most popular deep neural networks. The network was introduced in 2015 by He et al [21]. ResNet has been one of the most successful and popular deep learning models to date. ResNet won the 2015 ILSVRC challenge. ResNet allows you to train more than 150 layers. Before ResNet, deep neural networks faced the problem of vanishing gradient or exploding gradient if the layers increased. But ResNet was able to overcome this problem by using skip connection. For ResNet, 18-layer, 50-layer, 101-layer, and 150-layer versions are available, known as ResNet18, ResNet50, ResNet101, and ResNet152 breads, respectively. ResNet18 is used as a backbone in the Siamese network in this article.

By introducing a model called ViT (Vision Transformer), Dosovitskiy et al. [22] achieved excellent results on a large private-labeled image dataset without the use of convolution layers. By adding distillation token to the ViT model, Touvron et al. [23] introduced the DeiT (data-efficient transformers) model. Their model on the ImageNet database was able to find competitive results with the ViT model. The ViT and DeiT models are based on multiheaded self-attention. the output of an attention block is as follows [24]:

$$Attention(Q, K, V) = softmax(\frac{QK^T}{\sqrt{d_k}})V, \tag{6}$$

where Q represents the queries matrix, K represents the keys matrix, and V represents the values matrix. There are different versions of the DeiT model that differ in the number of parameters and the model's size. In this article, the Tiny DeiT version is used as a backbone in the Siamese network.

IV. Proposed weight sharing approaches

Weight sharing in Siamese networks means that the network branches share their weights [15]. The question is whether network performance will be affected if there is no weight sharing. We assume that a Siamese structure is made of two networks (i.e., two branches), and each branch is composed of two parts: fully connected (FC) and CNN. In this study, the effects of weight sharing in Siamese networks have been investigated in four cases: 1- Full weight sharing (FWS): FC and CNN weights are fully shared. 2- Partial weight sharing (PWS) in FC Weights are shared only in the FC section, and there is no weight sharing in the CNN part. 3- Partial weight sharing (PWS) in CNN: Weights are shared only on CNN, and there is no weight sharing in the FC part. 4- Without weight sharing (WWS): No weight is shared between the branches of the Siamese network. It should be noted that the SqueezeNet, GoogleNet, ResNet18, and Tiny DeiT backbones have been used to implement the CNN section of the networks. Three

linear layers with ReLU activation function have also been used to implement FC. The embedding space is a vector with dimensions of 500.

V. Experimental results

The results were evaluated in the following four cases. For a detailed analysis of the results, radar diagrams of three measures for different backbones are drawn in Figures 1 to 4. In these diagrams, if the chart is larger than others and includes others, it means that the results, in that case, are better than the others. The diagrams are based on the data in Tables I, II, III, and IV.

A. Contrastive loss on the AT&T face database

Table I shows the results of different backbones on the face database and contrastive loss. According to the results, in all cases except Tiny DeiT, the accuracy in "PWS in FC" is better than all other cases, and in Tiny DeiT, the best accuracy is 0.88 for both the "PWS in FC" and "FWS" cases. The radar diagram for the results of this mode is drawn in Figure 1. . In most cases, the triangle area created for "PWS in FC" is larger than that of "FWS," which indicates better network performance in this case. The "PWS in CNN" and "WWS" charts in most cases had poorer results than "PWS in FC."

TABLE I: Results on the face database with contrastive loss

		WWS	PWS in CNN	PWS in FC	FWS
SqueezeNet	Accuracy	0.55	0.51	**0.91**	0.90
	Precision	0.49	0.44	**0.91**	0.86
	Recall	0.57	0.43	0.89	**0.91**
GoogleNet	Accuracy	0.85	0.84	**0.91**	0.90
	Precision	0.84	0.87	**0.97**	0.86
	Recall	0.88	0.81	0.86	**0.91**
ResNet18	Accuracy	0.89	0.90	**0.93**	0.90
	Precision	**1.00**	0.95	0.89	0.86
	Recall	0.80	0.87	**0.94**	0.91
Tiny DeiT	Accuracy	0.44	0.44	**0.88**	**0.88**
	Precision	0.44	0.44	0.93	**0.96**
	Recall	**1.00**	**1.00**	0.77	0.74

B. Contrastive loss on Fingerprint dataset

Table II shows the results of different backbones on the fingerprint dataset and contrastive loss. The noticeable point about this table is the accuracy of ResNet18. The accuracy is better than "FWS" even on "PWS in CNN." As seen in Figure 2, in most cases, the areas of triangles for the "PWS in FC" mode are larger than the other cases. In general, it can be concluded that the best accuracy is obtained in "PWS in FC." The results are the best when using ResNet18 in both "PWS in FC" and "PWS in CNN" modes. Also, the highest precision and recall are for "WWS" and "FWS," respectively. The important thing is to get the highest possible accuracy using GoogleNet in both the "PWS in FC" and "FWS" modes.

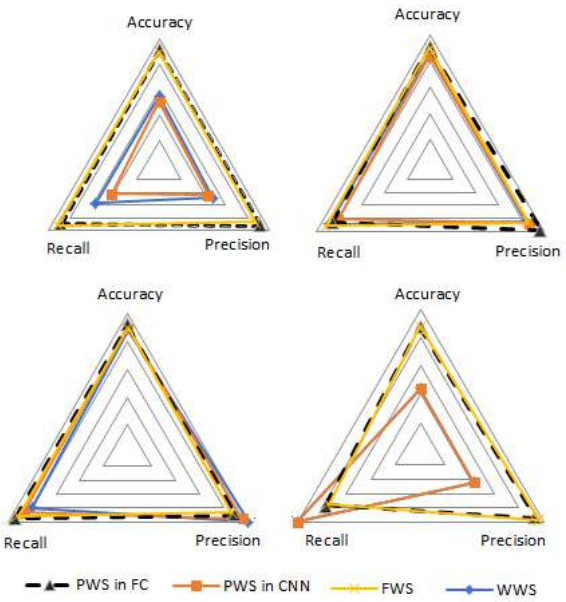


Fig. 1: Radar charts of measures for the face database with contrastive loss. Top left: SqueezeNet. Top right: GoogleNet. Bottom left: ResNet18. Bottom right: Tiny DeiT.

TABLE II: Results on fingerprint dataset with contrastive loss

		WWS	PWS in CNN	PWS in FC	FWS
SqueezeNet	Accuracy	0.53	0.46	**0.87**	0.84
	Precision	0.50	0.45	0.90	**1.00**
	Recall	**1.00**	0.74	0.81	0.66
GoogleNet	Accuracy	0.93	0.87	**1.00**	**1.00**
	Precision	0.95	0.86	**1.00**	**1.00**
	Recall	0.88	0.84	**1.00**	**1.00**
ResNet18	Accuracy	0.91	**0.94**	**0.94**	0.92
	Precision	**0.96**	0.94	0.93	0.88
	Recall	0.87	0.94	0.96	**0.98**
Tiny DeiT	Accuracy	0.50	0.50	**0.88**	0.79
	Precision	0.38	0.42	**0.91**	**0.91**
	Recall	0.11	0.17	**0.83**	0.62

C. Proxy Anchor loss on AT&T face database

Table III shows the results of different backbones on the face database and proxy anchor loss. On SqueezeNet and ResNet18 backbones, "PWS in FC" accuracy is higher than "FWS." The GoogleNet and Tiny DeiT backbones in "FWS" have the best accuracy. Figure 3 illustrates a general comparison in terms of three performance measures. The areas of the SqueezeNet and ResNet18 triangles in "PWS in FC" mode are the largest. On GoogleNet and Tiny DeiT, the triangles in "PWS in FC" and "FWS" are approximately the same.

D. Proxy Anchor loss on Fingerprint dataset

Table IV shows the results of different backbones on the fingerprint dataset and proxy anchor Loss. On SqueezeNet and ReSnet18 backbones, the accuracy of "FWS" is better than

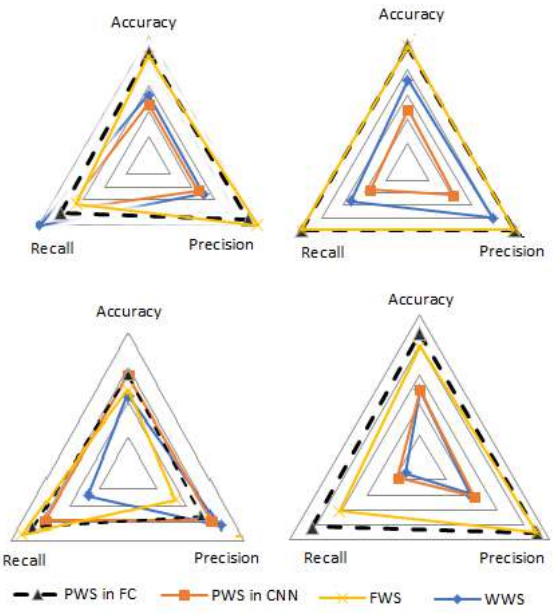


Fig. 2: Radar charts of measures for the fingerprint dataset with contrastive loss. Top left: SqueezeNet. Top right: GoogleNet. Bottom left: ResNet18. Bottom right: Tiny DeiT.

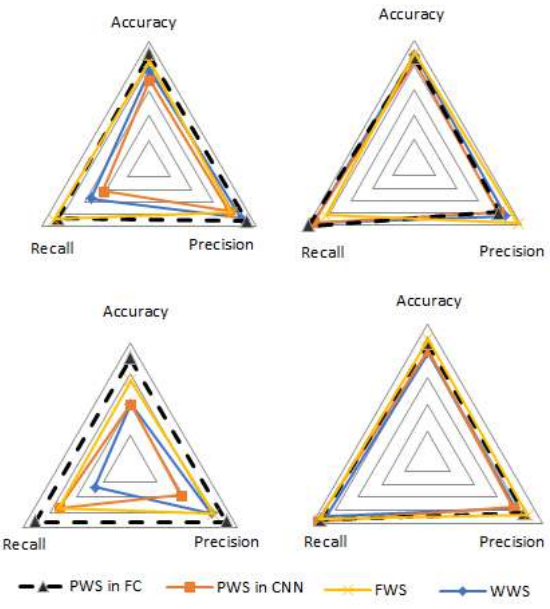


Fig. 3: Radar charts of measures for the face database with proxy anchor loss. Top left: SqueezeNet. Top right: GoogleNet. Bottom left: ResNet18. Bottom right: Tiny DeiT.

TABLE III: Results on the face database with proxy anchor loss

		WWS	PWS in CNN	PWS in FC	FWS
SqueezeNet	Accuracy	0.76	0.69	**0.90**	0.84
	Precision	0.86	0.75	**0.91**	0.78
	Recall	0.54	0.43	0.86	**0.89**
GoogleNet	Accuracy	0.89	0.84	0.86	**0.90**
	Precision	0.85	0.78	0.79	**0.97**
	Recall	0.95	0.95	**1.00**	0.83
ResNet18	Accuracy	0.90	0.90	**0.98**	0.94
	Precision	0.95	0.89	**0.98**	0.95
	Recall	0.87	0.93	**0.98**	0.93
Tiny DeiT	Accuracy	0.79	0.80	0.86	**0.89**
	Precision	0.76	0.74	0.83	**0.87**
	Recall	0.88	**0.95**	0.93	0.93

TABLE IV: Results on the fingerprint dataset with proxy anchor loss

		WWS	PWS in CNN	PWS in FC	FWS
SqueezeNet	Accuracy	0.69	0.64	0.69	**0.72**
	Precision	0.67	0.63	0.65	**0.69**
	Recall	0.68	0.57	0.72	**0.74**
GoogleNet	Accuracy	0.85	0.80	**0.90**	0.87
	Precision	**0.89**	0.71	0.88	0.83
	Recall	0.74	**0.91**	0.88	0.88
ResNet18	Accuracy	0.79	0.83	0.82	**0.87**
	Precision	0.82	**0.89**	0.86	0.83
	Recall	0.77	0.77	0.79	**0.94**
Tiny DeiT	Accuracy	0.68	0.71	**0.82**	0.78
	Precision	0.65	0.82	**0.84**	0.80
	Recall	0.68	0.42	**0.72**	0.65

other modes. The GoogleNet and Tiny DeiT backbones have the best accuracy for the ”PWS in CNN” modes. As seen in Figure 4 for Tiny DeiT the ”PWS in FC” triangle has the largest area. In SqueezeNet and GoogleNet, the ”PWS in FC” and ”FWS” mode triangles have approximately equal areas. In ResNet18, the ”FWS” mode has the largest triangle.

VI. Conclusion

This paper investigates the effects of weight sharing in four different modes on Siamese networks. Contrastive loss and proxy anchor loss functions were used for this purpose. The models were trained on the AT&T face database and the fingerprint dataset. According to the results, in most cases, the accuracy of contrastive loss in ”partial weight sharing in the FC” mode was better than ”full weight sharing.”

References

[1] G. Koch, R. Zemel, R. Salakhutdinov, et al., Siamese neural networks for one-shot image recognition, in: ICML deep learning workshop, Vol. 2, Lille, 2015.

[2] D. Chicco, Siamese neural networks: An overview, Artificial Neural Networks (2021) 73–94.

[3] B. Ghojogh, M. Sikaroudi, S. Shafiei, H. R. Tizhoosh, F. Karray, M. Crowley, Fisher discriminant triplet and contrastive losses for training siamese networks, in: 2020 international joint conference on neural networks (IJCNN), IEEE, 2020, pp. 1–7.

[4] I. Melekhov, J. Kannala, E. Rahtu, Siamese network features for image matching, in: 2016 23rd International Conference on Pattern Recognition (ICPR), IEEE, 2016, pp. 378–383.

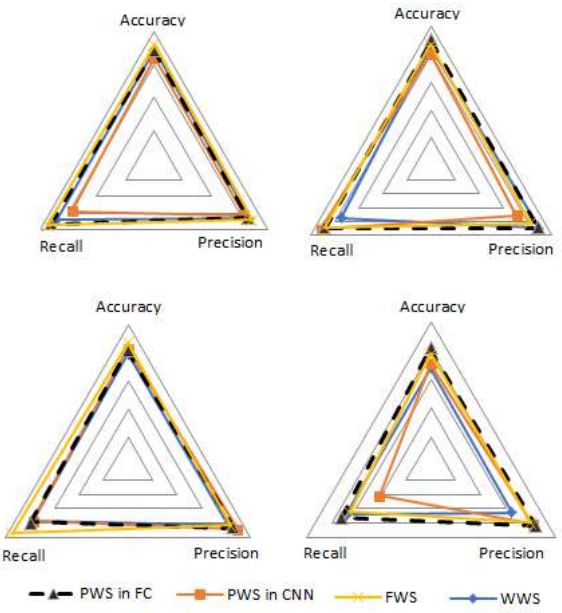


Fig. 4: Radar charts of measures for the fingerprint dataset with proxy anchor loss. Top left: SqueezeNet. Top right: GoogleNet. Bottom left: ResNet18. Bottom right: Tiny DeiT.

[5] J. Yu, M. Li, X. Hao, G. Xie, Deep fusion siamese network for automatic kinship verification, in: 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020), IEEE, 2020, pp. 892–899.

[6] C. Han, S. Shan, M. Kan, S. Wu, X. Chen, Face recognition with contrastive convolution, in: Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 118–134.

[7] X. Dong, J. Shen, Triplet loss in siamese network for object tracking, in: Proceedings of the European conference on computer vision (ECCV), 2018, pp. 459–474.

[8] L. Yang, M. Zhang, C. Li, M. Bendersky, M. Najork, Beyond 512 tokens: Siamese multi-depth transformer-based hierarchical encoder for long-form document matching, in: Proceedings of the 29th ACM International Conference on Information & Knowledge Management, 2020, pp. 1725–1734.

[9] Z. Zhang, H. Peng, Deeper and wider siamese networks for real-time visual tracking, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 4591–4600.

[10] Y. Cheng, H. Wang, A modified contrastive loss method for face recognition, Pattern Recognition Letters 125 (2019) 785–790.

[11] Z. He, W. Su, Z. Bi, M. Wei, Y. Dong, G. Xu, The improved siamese network in face recognition, in: 2019 International Conference on Intelligent Computing, Automation and Systems (ICICAS), IEEE, 2019, pp. 443–446.

[12] S. Ghosh, S. Ghosh, P. Kumar, E. Scheme, P. P. Roy, A novel spatio-temporal siamese network for 3d signature recognition, Pattern Recognition Letters 144 (2021) 13–20.

[13] D. D. Chakladar, P. Kumar, P. P. Roy, D. P. Dogra, E. Scheme, V. Chang, A multimodal-siamese neural network (msnn) for person verification using signatures and eeg, Information Fusion 71 (2021) 17–27.

[14] L. Lu, Y. Su, G. E. Karniadakis, Collapse of deep and narrow neural nets, arXiv preprint arXiv:1808.04947 (2018).

[15] X. Chen, K. He, Exploring simple siamese representation learning, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 15750–15758.

[16] T. Chen, S. Kornblith, M. Norouzi, G. Hinton, A simple framework for contrastive learning of visual representations, in: International conference on machine learning, PMLR, 2020, pp. 1597–1607.

[17] S. Kim, D. Kim, M. Cho, S. Kwak, Proxy anchor loss for deep metric learning, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 3238–3247.

[18] R. Hadsell, S. Chopra, Y. LeCun, Dimensionality reduction by learning an invariant mapping, in: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), Vol. 2, IEEE, 2006, pp. 1735–1742.

[19] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, K. Keutzer, Squeezenet: Alexnet-level accuracy with 50x fewer parameters and¡ 0.5 mb model size, arXiv preprint arXiv:1602.07360 (2016).

[20] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 1–9.

[21] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

[22] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, et al., An image is worth 16x16 words: Transformers for image recognition at scale, arXiv preprint arXiv:2010.11929 (2020).

[23] H. Touvron, M. Cord, M. Douze, F. Massa, A. Sablayrolles, H. Jégou, Training data-efficient image transformers & distillation through attention, in: International Conference on Machine Learning, PMLR, 2021, pp. 10347–10357.

[24] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, I. Polosukhin, Attention is all you need, in: Advances in neural information processing systems, 2017, pp. 5998–6008.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Hyper Ideals in Hyper Equality Algebras

1st Akbar Paad, 2nd Mahmood Bakhshi

1,2nd Department of Mathematics, *University of Bojnord*, Bojnord, Iran, akbar.paad@gmail.com, bakhshi@ub.ac.ir

***Abstract*—In this paper, the concept of (strong) hyper ideals in bounded hyper equality algebras which is a generalization of ideals in bounded equality algebras are introduced. Also, the relations between strong hyper ideals and hyper deductive systems in good bounded hyper equality algebras are studied and in the follow, the notion of involutive hyper equality algebras are introduced. Moreover, A regular hyper congruence relation via a strong hyper ideal in good involutive hyper equality algebras are constructed and finally, the quotient hyper equality algebras via hyper strong hyper ideals are provided.**

***Index Terms*— (Involutive) Hyper equality algebra, (strong, weak) hyper ideal, hyper congruence.**

I. Introduction

Fuzzy type theory (FTT) has been developed by Novák as a fuzzy logic of higher order, the fuzzy version of the classical type theory of the classical logic of higher order. EQ-algebras generalize the residuated lattices that have three binary operations meet, multiplication, fuzzy equality and a unit element. If the product operation in EQ-algebras is replaced by another binary operation smaller or equal than the original product we still obtain an EQ-algebra, and this fact might make it difficult to obtain certain algebraic results. For this reason, equality algebras were introduced by Jeni [10], which the motivation cames from EQ-algebras [13]. The concept of ideals in bounded equality algebras were introduced by Paad [14], which the motivation cames from BL-algebras. Hyper structure theory was introduced by Marty [12], at the 8th Congress of Scandinavian Mathematicians. Till now, the hyper structures are studied in many subject of pure and applied mathematics. For example hyper structure theory has been intensively researched in [1, 2, 6, 8, 11, 15]. Recently, Borzooei has applied hyper theory to EQ-algebras to introduce hyper EQ-algebras [3] which are a generalization of EQ-algebras and Cheng et. al [4], has applied hyper theory to equality algebras to introduce hyper equality algebras which are a generalization of equality algebras. Hyper filters (or hyper deductive systems) are important tools in studying hyper structures [1, 5]. The above are the motivation of introducing and studying (strong)hyper ideals in bounded hyper equality algebras which is a generalization of ideals in bounded equality algebras.

II. Preliminaries

Definition 1. *A hyper equality algebra is a nonempty set H endowed with a binary operation $\wedge$, a binary hyperoperation $\sim$ satisfying the following axioms, for all $x, y, z \in H$:*

$(HE1)$ $(H; \wedge)$ *is a meet-semilattice with top element 1 (with the order $x \leq y \Leftrightarrow x \wedge y = x$);*

$(HE2)$ $x \sim y \ll y \sim x$;

$(HE3)$ $1 \in x \sim x$;

$(HE4)$ $x \in 1 \sim x$;

$(HE5)$ $x \leq y \leq z$ *implies* $x \sim z \ll y \sim z$ *and* $x \sim z \ll x \sim y$;

$(HE6)$ $x \sim y \ll (x \wedge z) \sim (y \wedge z)$;

$(HE7)$ $x \sim y \ll (x \sim z) \sim (y \sim z)$.

where $W \ll Z$ is defined by for all $w \in W$, there exists $z \in Z$ such that $w \leq z$.

The auxiliry hyperoperation $\rightarrow$ is defined as $x \rightarrow y := x \sim (x \wedge y)$ and $\leftrightarrow$ is defined as $x \leftrightarrow y := (x \rightarrow y) \wedge (y \rightarrow x)$ for all $x, y \in H$. Moreover, for any nonempty subsets $A, B \subseteq H$, we write $A \wedge B = \{a \wedge b \mid a \in A, b \in B\}$ and $A \circ B = \bigcup_{a \in A, b \in B} a \circ b$, where $\circ \in \{\sim, \rightarrow, \leftrightarrow\}$. Hyper equality algebra H is said to be *bounded* if it has a bottom element 0 (with respect the order $\leq$). When H is bounded the set $x \rightarrow 0 = x \sim 0$ is denoted by $\neg x$ and H is called good if $x = 1 \sim x$, for all $x \in H$. Moreover, H is called symmetric if $x \sim y = y \sim x$, for all $x, y \in H$.

Definition 2. *A nonempty subset S containing 1 of hyper equality algebra H is called a hyper subalgebra, if S is a hyper equality algebra with respect to the hyper operation $\sim$ and the binary operation $\wedge$ on H.*

Definition 3. *[14] Let H be a bounded equality algebra and I a nonempty subset of H. Then I is called an ideal of H, if it satisfies:*
(i) if $x \leq y$ and $y \in I$, then $x \in I$, for every $x, y \in E$,
(ii) for every $x, y \in I$, $x^- \rightarrow y \in I$.

Definition 4. *A nonempty subset S containing 1 of hyper equality algebra H is said to be a hyper subalgebra, if S is*

a hyper equality algebra with respect to the hyper operation $\sim$ and the binary operation $\wedge$ on H.

Definition 5. *[14] Let H be a bounded equality algebra and I a nonempty subset of H. Then I is called an ideal of H, if it satisfies:*
(i) if $x \leq y$ and $y \in I$, then $x \in I$, for every $x, y \in E$,
(ii) for every $x, y \in I$, $x^- \to y \in I$.

Definition 6. *Let F be a nonempty subset of hyper equality algebra H, satisfying*
(F) $x \in F$ and $x \leq y$ imply $y \in F$ for all $x, y \in H$. Then F is called a

- weak hyper filter, *if*
  (WHF) $x \in F$ and $x \sim y \subseteq F$ imply $y \in F$, for all $x, y \in H$.
- hyper filter, *if*
  (HF) $x \in F$ and $F \ll x \sim y$ imply $y \in F$, for all $x, y \in H$.
- strong hyper filter, *if*
  (SHF) $x \in F$ and $x \sim y \cap F \neq \emptyset$ imply $y \in F$, for all $x, y \in H$.

Definition 7. *Let D be a nonempty subset of hyper equality algebra H. Then D is called a*

- weak hyper deductive system, *if D satisfies (F) and*
  (WHD) $x \in D$ and $x \to y \subseteq D$ imply $y \in D$, for all $x, y \in H$.
- hyper deductive system, *if D satisfies (F) and*
  (HD) $x \in D$ and $D \ll x \to y$ imply $y \in D$, for all $x, y \in H$.
- strong hyper deductive system, *if $1 \in D$ and*
  (SHD) $x \in D$ and $x \to y \cap D \neq \emptyset$ imply $y \in D$, for all $x, y \in H$.

Definition 8. *Let θ be an equivalence relation on hyper equality algebra H.*
(1) For any $A, B \subseteq H$, $A\bar{\theta}B$ means for any $a \in A$ there exists $b \in B$ such that $a\theta b$ and for any $b \in B$ there exists $a \in A$ such that $a\theta b$.
(2) For any $A, B \subseteq H$, $A\bar{\bar{\theta}}B$ means for any $a \in A$ and any $b \in B$ such that $a\theta b$.
(3) θ is called a hyper congruence relation if for all $x, y, u, v \in H$, $x\theta y$ and $u\theta v$ imply $x \sim u\bar{\theta}y \sim v$, $x \wedge u\theta y \wedge v$.
(4) θ is called a strong hyper congruence relation if for all $x, y, u, v \in H$, $x\theta y$ and $u\theta v$ imply $x \sim u\bar{\bar{\theta}}y \sim v$, $x \wedge u\theta y \wedge v$. Not that every strong hyper congruence relation is a hyper congruence relation.

Definition 9. *A hyper congruence relation θ on H is called regular if for all $x, y \in H$, $[x]_\theta \leq_\theta [y]_\theta$ implies $[x]_\theta \ll [y]_\theta$.*

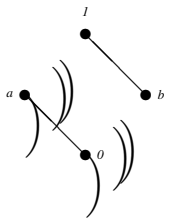
III. Hyper Ideals in Bounded Hyper Equality Algebras

Definition 10. *Let I be a nonempty downer subset of H;*
$(HI1)$ *$x \leq y$ and $y \in I$ imply $x \in I$, for all $x, y \in H$.*
I is called a

- weak hyper ideal *if $x, y \in I$ imply $\neg x \to y \subseteq I$, for all $x, y \in H$.*
- hyper ideal *if $x, y \in I$ imply $\neg x \to y \ll I$, for all $x, y \in H$.*
- strong hyper ideal *if $x, y \in I$ imply $(\neg x \to y) \cap I \neq \emptyset$, for all $x, y \in H$.*

Denote $HId(H)$ and $SHId(H)$ the set of all hyper ideals and the set of strong hyper ideals of a bounded hyper equality algebra H, respectively.

Example 1. *[4] Let $H = \{0, a, b, 1\}$ in which the Hasse diagram and operations $\sim$ and $\wedge$ on H are as follows:*



$\wedge$	0	a	b	1
0	0	0	0	0
a	0	a	0	a
b	0	0	b	b
1	0	a	b	1

$\sim$	0	a	b	1
0	$\{1\}$	$\{1\}$	$\{b, 1\}$	$\{0, a\}$
a	$\{1\}$	$\{1\}$	$\{a, 1\}$	$\{a\}$
b	$\{b, 1\}$	$\{a, 1\}$	$\{1\}$	$\{b, 1\}$
1	$\{0, a\}$	$\{a\}$	$\{b, 1\}$	$\{1\}$

Then $(H, \wedge, \sim, 0, 1)$ is a bounded hyper equality algebra and $I = \{0, a\}$ is a hyper ideal of H and $J = \{0, a, b\}$ is a strong hyper ideal of H which is not hyper ideal. Because, $a, b \in J$ and $\neg b \to a = \{a, b, 1\} \nsubseteq J$.

Theorem 1. *In any bounded hyper equality algebra,*
(i) Hyper ideals and weak hyper ideals are coincide.
(ii) Every (weak)hyper ideal is a strong hyper ideal.

Proof. (i) Let I be a hyper ideal of H and $x, y \in I$. Then $\neg x \to y \ll I$. If $a \in \neg x \to y$, then there exists $b \in I$, such that $a \leq b$. Hence, by $(HI1)$, $a \in I$ and so $\neg x \to y \subseteq I$. Therefore, I is a weak hyper ideal of H. Conversely, let I be a weak hyper ideal of H and $x, y \in I$. Then $\neg x \to y \subseteq I$ and so for any $a \in \neg x \to y$, $a \in I$ and since $a \leq a$, we get that $\neg x \to y \ll I$. Therefore, I is a hyper ideal of H.
(ii) Let I be a hyper ideal of H and $x, y \in I$. Then $\neg x \to y \ll I$. By $(HE8)$, $y \ll \neg x \to y$ and so there exists $b \in \neg x \to y$, such that $y \leq b$ and since $\neg x \to y \ll I$, so there exists $c \in I$, such that $b \leq c$. Now, by $(HI1)$, we conclude that $b \in I$ and so $y \in I$. Hence, $y \in (\neg x \to y) \cap I$ and so $(\neg x \to y) \cap I \neq \emptyset$. Therefore, I is a strong hyper ideal of H. Moreover, by (i), I is a weak hyper ideal and so I is a strong hyper ideal of H.

Definition 11. *Let D be a nonempty subset of bounded hyper equality algebra H. Then*
(i) D is called S-reflexive, if $\neg A \cap D \neq \emptyset$, implies $\neg A \subseteq D$, for any $A \subseteq H$.
(ii) D is called $S_\neg$-reflexive, if $(\neg x \to y) \cap D \neq \emptyset$, implies $(\neg x \to y) \subseteq D$, for any $x, y \in H$.

Example 2. *[4] Let $H = \{0, a, 1\}$ with $0 < a < 1$. Define operations $\wedge$, and hyper operation $\sim$ as follows: $x \wedge y = \min\{x, y\}$ and*

$\sim$	0	a	1
0	$\{1\}$	$\{0\}$	$\{0\}$
a	$\{0\}$	$\{1\}$	$\{a, 1\}$
1	$\{0\}$	$\{a, 1\}$	$\{1\}$

Then $(H, \wedge, \sim, 0, 1)$ is a bounded hyper equality algebra and $D = \{0, 1\}$ is S-reflexive subset of H. Moreover, D is $S_\neg$-reflexive subset of H.

Theorem 2. *Every $S_\neg$-reflexive strong hyper ideal of H, is a (weak) hyper ideal of H.*

Proof. The proof is straightforward.

By example 1, the converse of Theorem 2 is not correct in general.

Theorem 3. *Let I be a S-reflexive nonempty subset of good bounded hyper equality algebra H. Then I is a S-reflexive strong hyper ideal of H if and only if*
(i) $0 \in I$,
(ii) If $x \in I$ and $\neg(\neg x \to \neg y) \cap I \neq \emptyset$, then $y \in I$, for any $x, y \in H$.

Proof. Let I is a S-refexive strong hyper ideal of H. Since I is a nonempty, so there exists $x \in I$ and so by $0 \leq x$ and $(HI1)$, we get that $0 \in I$. Now, let $x \in I$ and $\neg(\neg x \to \neg y) \cap I \neq \emptyset$, for $x, y \in H$. Then $\neg(\neg x \to \neg y) \subseteq I$. We have $\neg x \to \neg y = \neg x \to (y \to 0) \ll y \to (\neg x \to 0) = y \to \neg\neg x$ and so we get that $y \ll (\neg x \to \neg y) \to \neg\neg x$. Moreover, we have $(\neg x \to \neg y) \to \neg\neg x = (\neg x \to \neg y) \to (\neg x \to 0) \ll \neg x \to \neg(\neg x \to \neg y)$ Hence, $y \ll \neg x \to \neg(\neg x \to \neg y)$. Now, since $\{x\} \subseteq I$ and $\neg(\neg x \to \neg y) \subseteq I$, by we get that

$$\neg x \to \neg(\neg x \to \neg y) \subseteq I$$

By $y \ll \neg x \to \neg(\neg x \to \neg y)$, there exists $b \in \neg x \to \neg(\neg x \to \neg y) \subseteq I$, such that $y \leq b$ and so by $(HI1)$, we get that $y \in I$.

Conversely, let $x \leq y$ and $y \in I$, for $x, y \in H$. Then by $(HE10)$, $\neg y \ll \neg x$ and so $1 \in \neg y \to \neg x$ and so $0 \in \neg 1 \subseteq \neg y \to \neg x$ and since $0 \in I$, we get that $0 \in \neg(\neg y \to \neg x) \cap I$. Hence, $\neg(\neg y \to \neg x) \cap I \neq \emptyset$ and since $y \in I$, by (ii) we conclude that $x \in I$. Therefore, I satisfies in $(HI1)$. Let $x, y \in I$. Then we have $(\neg x \to y) \to y \ll (y \to 0) \to ((\neg x \to y) \to 0) = \neg y \to \neg(\neg x \to y)$ And so by there exists $b \in \neg y \to \neg(\neg x \to y)$, such that

$$\neg b \ll \neg((\neg x \to y) \to y)$$

Since $\neg x \ll (\neg x \to y) \to y$, hence there exists $w \in (\neg x \to y) \to y$ such that $\neg w \ll \neg\neg x$ and since by we have $\neg x \ll (\neg x \to 0) \to 0 = \neg\neg\neg x$, we conclude that $1 \in (\neg x \to \neg\neg\neg x)$ and so $0 \in \neg 1 \subseteq \neg(\neg x \to \neg\neg\neg x)$ and since $0 \in I$, we get that $\neg(\neg x \to \neg\neg\neg x) \cap I \neq \emptyset$ and since $x \in I$, we get that $\{x\} \subseteq I$. Hence, by $\neg\neg x \cap I \neq \emptyset$ and so $\neg\neg x \subseteq I$. Now, since $\neg w \ll \neg\neg x$, by $(HI1)$, we conclude that $\neg w \cap I \neq \emptyset$ and since $w \in (\neg x \to y) \to y$, we get that $\neg w \subseteq \neg((\neg x \to y) \to y)$. Hence, $\neg((\neg x \to y) \to y) \cap I \neq \emptyset$ and so $\neg((\neg x \to y) \to y) \subseteq I$. Now, since $\neg b \ll \neg((\neg x \to y) \to y)$, by $(HI1)$, we conclude that $\neg b \cap I \neq \emptyset$ and since $b \in \neg y \to \neg(\neg x \to y)$, we get that $\neg b \subseteq \neg(\neg y \to \neg(\neg x \to y))$. Hence, by $(HI1)$, we conclude that $\neg(\neg y \to \neg(\neg x \to y)) \cap I \neq \emptyset$ and since $y \in I$, we get that $\{y\} \subseteq I$. Hence, $(\neg x \to y) \cap I \neq \emptyset$. Therefore, I is a S-reflexive (strong, weak)hyper ideal of H.

Theorem 4. *Let F be a S-reflexive strong hyper deductive system of good bounded hyper equality algebra H. Then $N(F)$ is a strong hyper ideal of H.*

Proof. The proof is straightforward.

Theorem 5. *Let I be a S-reflexive strong hyper ideal of good bounded hyper equality algebra H. Then $N(I)$ is a strong hyper deductive system of H.*

Proof. The proof is straightforward.

There is open problem how can reach the converse of Theorems 4 and 5?

Theorem 6. *Let $\{I_\alpha\}_{\alpha \in A}$ be a family of (weak)hyper ideal of H. Then $\bigcap_{\alpha \in A} I_\alpha$ is a (weak)hyper ideal of H.*

Proof. The proof is straightforward.

Theorem 7. *Let $\{I_\alpha\}_{\alpha \in A}$ be a family of $S_\neg$-reflexive strong hyper ideal of H. Then $\bigcap_{\alpha \in A} I_\alpha$ is a $S_\neg$-reflexive strong hyper ideal of H.*

Proof. The proof is straightforward.

Theorem 8. *Let $\{I_\alpha\}_{\alpha \in A}$ be a family of ($S_\neg$-reflexive strong) hyper ideal of H. Then*
(i) If $\{I_\alpha\}_{\alpha \in A}$ is a chain, then $\bigcup_{\alpha \in A} I_\alpha$ is a ($S_\neg$-reflexive strong) hyper ideal of H.
(ii) If H is a chain, then $\bigcup_{\alpha \in A} I_\alpha$ is a ($S_\neg$-reflexive strong) hyper ideal of H.

Proof. The proof is straightforward.

Definition 12. *A bounded hyper equality algebra $(H; \sim, \wedge, 1)$ is called involutive, if $x \in \neg\neg x$, for all $x \in H$.*

Example 3. *[4] Let $H = \{0, a, 1\}$. Define operations $\wedge$ and $\sim$ on H as follows: $x \wedge y = \min\{x, y\}$ and*

$\sim$	0	a	1
0	$\{1\}$	$\{0, a\}$	$\{0\}$
a	$\{0, a\}$	$\{1\}$	$\{a\}$
1	$\{0\}$	$\{a\}$	$\{1\}$

Then $(H; \sim, \wedge, 1)$ is an involutive hyper equality algebra.

Theorem 9. *Let I be a S-reflexive strong hyper ideal of symmetric good involutive hyper equality algebra H. Then the binary relation $\equiv_I$ on H which is defined by*

$$x \equiv_I y \text{ if and only if } \neg(\neg x \sim \neg y) \cap I \neq \emptyset$$

is a strong hyper congruence relation on H.

Proof. It is easy to check that $\equiv_I$ is an equivalence relation on H.

Now, we prove that equivalence relation $\equiv_I$ is compatible with $\wedge$. Let $x, y, z \in H$ and $x \equiv_I y$. Then $\neg(\neg x \sim \neg y) \cap I \neq \emptyset$ and so $\neg(\neg x \sim \neg y) \subseteq I$. Since H is an involutive hyper equality algebra, we get that $x \in \neg\neg x$ and $y \in \neg\neg y$ and so $x \sim y \subseteq \neg\neg x \sim \neg\neg y$. Hence, $\neg(x \sim y) \subseteq \neg(\neg\neg x \sim \neg\neg y)$. we have $(x \wedge z) \sim (y \wedge z) \ll \neg(x \wedge z) \sim \neg(y \wedge z)$ So there exists $b \in \neg(x \wedge z) \sim \neg(y \wedge z)$, such that

$$\neg b \ll \neg((x \wedge z) \sim (y \wedge z))$$

Since $x \sim y \ll (x \wedge z) \sim (y \wedge z)$, hence there exists $w \in (x \wedge z) \sim (y \wedge z)$, such that $\neg w \ll \neg(x \sim y)$ and since $\neg(x \sim y) \subseteq \neg(\neg\neg x \sim \neg\neg y)$, we get that $\neg w \ll \neg(\neg\neg x \sim \neg\neg y)$. Moreover, we have $\neg x \sim \neg y \ll \neg\neg x \sim \neg\neg y$. So there exists $p \in \neg\neg x \sim \neg\neg y$, such that $\neg p \ll \neg(\neg x \sim \neg y)$ and since $\neg(\neg x \sim \neg y) \subseteq I$, we get that $\neg p \ll I$ and so by $(HI1)$, $\neg p \cap I \neq \emptyset$. Now, since $p \in \neg\neg x \sim \neg\neg y$, we get that $\neg p \subseteq \neg(\neg\neg x \sim \neg\neg y)$. Hence, $\neg(\neg\neg x \sim \neg\neg y) \cap I \neq \emptyset$ and so $\neg(\neg\neg x \sim \neg\neg y) \subseteq I$ and by $\neg w \ll \neg(\neg\neg x \sim \neg\neg y)$, we conclude that $\neg w \ll I$ and so by $(HI1)$, $\neg w \cap I \neq \emptyset$. Now, since $w \in (x \wedge z) \sim (y \wedge z)$, we get that $\neg w \subseteq \neg((x \wedge z) \sim (y \wedge z))$. Hence, $\neg((x \wedge z) \sim (y \wedge z)) \cap I \neq \emptyset$ and so $\neg((x \wedge z) \sim (y \wedge z)) \subseteq I$. By $\neg b \ll \neg((x \wedge z) \sim (y \wedge z))$ and $\neg((x \wedge z) \sim (y \wedge z)) \subseteq I$, we conclude that $\neg b \ll I$ and so by $(HI1)$, $\neg b \cap I \neq \emptyset$ and since $b \in \neg(x \wedge z) \sim \neg(y \wedge z)$, we get that $\neg b \subseteq \neg(\neg(x \wedge z) \sim \neg(y \wedge z))$. Hence,

$$\neg(\neg(x \wedge z) \sim \neg(y \wedge z)) \cap I \neq \emptyset$$

Therefore, $x \wedge z \equiv_I y \wedge z$. Now, if $x \equiv_I y$ and $z \equiv_I w$, then $x \wedge z \equiv_I y \wedge z$ and $y \wedge z = z \wedge y \equiv_I w \wedge y = y \wedge w$ and so by transitivity we get that $x \wedge z \equiv_I y \wedge w$. By similar way $\equiv_I$ is compatible with $\sim$.

Therefore, $\equiv_I$ is a congruence relation on H. Finally, we show that $\equiv_I$ is a strong hyper congruence relation on H. Note that, if $x \equiv_I y$ and $z \equiv_I w$, then $x \sim z \equiv_I y \sim w$ and so $\neg(\neg(x \sim z) \sim \neg(y \sim w)) \cap I \neq \emptyset$. Hence, $\neg(\neg(x \sim z) \sim \neg(y \sim w)) \subseteq I$. Thus, for any $\alpha \in (x \sim z)$ and $\beta \in (y \sim w)$, we have $\neg(\neg\alpha \sim \neg\beta) \subseteq \neg(\neg(x \sim z) \sim \neg(y \sim w)) \subseteq I$ and so $\neg(\neg\alpha \sim \neg\beta) \cap I \neq \emptyset$. Hence, $\alpha \equiv_I \beta$ and so

$$x \sim z \overline{\equiv}_I y \sim w$$

Therefore, $\equiv_I$ is a strong hyper congruence relation on H.

Proposition 1. *Let $\frac{H}{\equiv_I} = \{[x] | x \in H\}$, where I is a S-reflexive strong hyper ideal of symmetric good involutive hyper equality algebra H. Define operation $\overline{\wedge}$ and hyper operation $\overline{\sim}$ on $\frac{H}{I}$ as follows:*

$$[x]\overline{\wedge}[y] := [x \wedge y], \quad [x]\overline{\sim}[y] = \{[t] | t \in x \sim y\}$$

Then $\overline{\wedge}$ and $\overline{\sim}$ are well defined.

Proof. It follows from Theorem 9.

Theorem 10. *Let I be a S-reflexive strong hyper ideal of chain symmetric good involutive hyper equality algebra H. Then $(\frac{H}{\equiv_I}; \overline{\sim}, \overline{\wedge}, [1])$ is a chain symmetric good involutive hyper equality algebra, which is called a quotient hyper equality algebra with respect to I.*

Proof. Let I be a S-reflexive strong hyper ideal of symmetric good involutive hyper equality algebra H. Then by Theorem 9, $\equiv_I$ is strong hyper congruence relation on H and since $\equiv_I$ is regular, we get that $(\frac{H}{\equiv_I}; \overline{\sim}, \overline{\wedge}, [1])$ is a hyper equality algebra and it is clear $\frac{H}{\equiv_I}$ is a chain. Now, we prove $\frac{H}{\equiv_I}$ is a symmetric good involutive hyper equality algebra. Let $[x], [y] \in \frac{H}{\equiv_I}$. Then $[x]\overline{\sim}[1] = [x \sim 1] = [x]$, so we get that $\frac{H}{\equiv_I}$ is a good and since H is a symmetric hyper equality algebra, we get that $[x]\overline{\sim}[y] = [x \sim y] = [y \sim x] = [y]\overline{\sim}[x]$ and so $\frac{H}{\equiv_I}$ is a symmetric hyper equality algebra. Moreover, we have $\neg\neg[x] = [\neg\neg x] = [\neg x \sim 0] = \{[t] \mid t \in \neg x \sim 0\} = \{[t] \mid t \in \neg\neg x\}$. Since H is an involutive, we get that $x \in \neg\neg x$ and so $[x] \in \neg\neg[x]$. Therefore, $\frac{H}{\equiv_I}$ is an ivolutive hyper equality algebra.

REFERENCES

[1] R. A. Borzooei, M. Bakhshi, O. Zahiri, "Filter theory on hyper residuated lattice," Quasigroups Relat. Syst. 22, pp. 33–50, 2014.

[2] R. A. Borzooei, A. Hasankhani, M. M. Zahedi, Y.B. Jun, "On hyper K-algebras," Math. Japonicae 52, pp. 113–121, 2000.

[3] R. A. Borzooei, B. G. Saffar, R. Ameri, " On hyper EQ-algebras," Italian J. Pure Appl. Mathematics-N. 31, pp. 77–96, 2013.

[4] X. Y. Cheng, X. L. Xin and Y. B. Jun, "Hyper equality algebras," Quantitative Logic and Soft Computing, Advances in Intelligent Systems and Computing 510, pp. 415–428, 2016.

[5] X. Y. Cheng, X. L. Xin, "Filter theory On hyper BE-algebras," Italian J. Pure Appl.Mathematics- N. 35, pp. pp. 509–526, 2015.

[6] L. C. Ciungu, "Internal states on equality algebras," Soft Computing, 19(4), pp. 939–953, 2015.

[7] F. Esteva, L. Godo, "Monoidal t-normbased logic: towards a logic for left-continuous t-norms," Fuzzy Sets and Systems, 124(3), pp. 271–288, 2001.

[8] S. Ghorbani, A. Hasankhani, E. Eslami, " Hyper MV-algebras," Set-Valued Math. Appl. 1, pp. 205–222, 2008.

[9] P. Hájek, "Metamathematics of Fuzzy Logic," Springer Science, 1998.

[10] S. Jenei, "Equality algebras, Studia Logica, 100(6), pp. 1201–1209, 2012.

[11] Y. B. Jun, M. M. Zahedi, X. L. Xin, R. A. Borzooei, " On hyper BCK-algebras," Italian J. Pure Appl. Mathematics-N 8, pp. 127–136, 2000.

[12] F. Marty, "Surune generalization de la notion de group," The 8th Congress Math. Scandinaves, Stockholm, pp. 45–49, 1934.

[13] V. Novák, B. De Baets, “EQ-algebras,” Fuzzy Sets and Systems, 160(20), pp. 2956–2978, 2009.
[14] A. Paad, “Ideals in bounded equality algebras,” Filomat 33 (7), pp. 2113–2123, 2019.
[15] X. L. Xin, “Hyper BCI-algebras,” Discuss. Math. General Algebra Appl. 26, pp. 5–19, 2006.
[16] F. Zebardast, R. A. Borzooei, M. Aaly Kologani, “Results on Equality algebras,” Information Sciences, 381, pp. 270–282, 2017.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Improved LOF Algorithm Using Random Point

1st Rasoul Sharify, 2nd Reza Heydari Gharaei and 3rd S. Mahmoud Taheri*

1st School of Engineering Science, College of Engineering, University of Tehran, Tehran, Iran, rasoul.sharifi@ut.ac.ir
2nd Department of Electrical Engineering, Shahid Bahonar University of Kerman, Kerman, Iran, reza.heydari@eng.uk.ac.ir
3rd School of Engineering Science, College of Engineering, University of Tehran, Tehran, Iran, sm_taheri@ut.ac.ir

***Abstract*— In the present study, Local outlier factor (LOF), the most popular and widely used density-based algorithm for outlier detection, is improved using a new idea. In the proposed idea, a random point in the neighborhood of all objects is used. This method improved the efficiency of the LOF algorithm to an acceptable level. Moreover, the results of the improved LOF indicated the competitiveness of the new algorithm in some real datasets. The comparisons are done based on two criteria: precision and AUC (Area Under the ROC Curve).**

***Index Terms*—outlier, outlier detection, random point, LOF algorithm, density-based methods**

I. Introduction

Outliers are referred to the data in a data set which are inconsistent with the rest of the data [1]. In other words, outliers are defined as data that are inconsistent with the expected model of data and deviate significantly from other data in such a way that it seems to be generated by a different procedure [2]. Outlier detection is the process of finding an object(s) in a dataset that do not conform to the normal behavior of objects [2]. These outliers can be points or clusters (groups of points).

Note:

Many outlier detection methods are basically the same despite having different names. Based on their applications, these methods are referred to by different names as follows: Anomaly Detection, Exception Detection, Novelty Detection, Error Detection, Fault Detection, Damage Detection, Peculiarity Detection, Noise Detection, Deviation Detection, and Outlier Detection.

Due to various applications of outlier detection in different fields, this issue is well received. Numerous applications of outlier detection show the significance and large scope of this issue. Fraud detection for credit cards [3], cell phones [4] and insurance [5], as well as, outlier detection in medical and public health [6], fault detection in safety-critical systems [7], web applications and data cleaning are examples of outlier detection applications [4]. For instance, an abnormal traffic pattern on a computer network could mean that a computer is hacked and is sending important information to an unknown destination. Moreover, an abnormal MRI image may be due to a malignant tumor. as another example, anomaly detection in a credit card transaction can indicate credit card fraud or identity theft [4].

One of the most famous outlier detection algorithms is $LocalOutlierFactoe(LOF)$ [8], which is classified as a nearest neighbor-based method (sub-category of relative density-based methods). Although several algorithms have been proposed to improve LOF over the years, it still has deficiencies in terms of speed and accuracy. Therefore, in the present study, LOF is improved using a random point in the neighborhood of objects. This random point is selected outside the dataset and then is considered as a neighborhood for all objects in the dataset Examining the results in some real datasets shows the efficiency of this innovative idea.

Note:

In general, a random point is not a neighbor for any of the data. But in proposed algorithm, despite the large distance from the random point to a data set, it is considered a neighbor for all the data.

In Section II, LOF algorithm will be explained in detail. In Section III, firstly, one of the shortcomings of LOF will be investigated. Then, the proposed idea to solve the problem will be explained, and finally, the improved LOF will be presented. In Section IV, first, some real datasets will be introduced; then, the results of the improved LOF will be compared with LOF algorithm, and the total results will be presented. In Section V, some conclusions will be presented and some recommendations will be made for fuerther research.

II. Literature Review

Inspired by [4] we review four subcategories for outlier detection methods as follows:

- Classification-Based methods
- Nearest neighbor-based methods
- Statistical methods
- Clustering-Based methods

A. Classification-Based outlier detection methods

An important sub-category of outlier detection algorithms is classification-based algorithms. The advantage of these methods is the speed of the test phase because they can quickly detect outliers after dataset classification. However, the disadvantage of these methods is their dependence on data labeling. Therefore, if accurate tags are not available, outlier detection becomes difficult [9].

B. Nearest neighbor-based outlier detection methods

In these methods the basis for identifying an outlier is the distance between data.

There are two main methods for detecting outliers based on distance. In the first method, the distance is globally assessed in the whole dataset. Considering the shortcoming of the first method, the second method measures the distance locally. The most significant outlier detection algorithm (LOF) belongs to this category, which measures the distance locally.

In the following, we describe two categories of nearest neighbor-based outlier detection methods.

1) The k-nearest neighbors (k-NN)-based methods: One outlier detection methods is the k-nearest neighbors algorithm (k-NN). In this method, the distance between an object and its neighbors is determined to define an outlier. These methods are among the oldest outlier detection methods. The shortcomings of distance-based methods are:

- The shortcoming of these methods is in detecting outliers in datasets with various densities.
- In these methods, the run time is long and increases significantly in objects with large volume and dimension.
- In neighborhood-based methods, measuring the distance between all objects to identify outliers causes many computational complexities.

To reduce the calculations in distance-based methods, reference points are used. Reference points are some fixed points in datasets by which relative neighborhood is measured. If there are a small number of these points, the precision of the algorithm decreases, but a large number of these points lead to the improvement of the calculations. Another solution to improve distance-based methods is using methods based on the relative density of data.

2) Relative density-based methods: The most notable outlier detection methods are density-based methods, which have been well-received in the last 2 decades. This study also presents an improved version of one of the most popular density-based methods (LOF) [8] .

C. Outlier detection using statistical methods

In outlier detection using statistical techniques, an outlier is referred to an observation that is irrelevant in terms of probability to other objects because it has not been generated by the assumed random model [10]. Statistical methods for detecting outliers are based on the key assumption that normal objects occur in high probability regions in a random model, while outliers occur in low probability ones. Statistical methods apply a statistical model to data (usually normal data). Then, a statistical test is used to check if a new object belongs to the statistical model. Those objects with a low probability of being generated by the learned model are detected as outliers.

Both parametric and non-parametric methods are employed to obtain a statistical model of outlier detection. However, parametric methods assume that the underlying distribution of objects is known, and distribution parameters are estimated from data. On the other hand, non-parametric methods assume that there is no knowledge of underlying distribution [11].

D. Clustering-Based outlier detection methods

In clustering, a dataset is divided into several clusters in such a way that data located in one cluster are similar to each other and different from the objects in other clusters. There are various methods for clustering objects, which are performed based on the type of objects, shape of clusters, and distance between objects. Several clustering-based outlier detection algorithms have been proposed so far such as KMOR [12] and DBSCAN [13] algorithms.

III. The improved LOF algorithm

In this section, we briefly introduce LOF algorithm first, then show the shortcomings of this algorithm in outlier detection in dense clusters, and finally explain our proposed algorithm for improving LOF algorithm.

A. LOF algorithm

Here we recall some basic definitions based on [8].

Definition 1 (k-th distance from an object)**.** *The k-th distance from an object ($k-distance(o)$) is the distance between an object and its k-th neighbor.*

Definition 2 (k neighborhood of an object)**.** *K nearest objects of a dataset to an object is called its k neighborhood:*

$$N_{k\text{-}distance(p)}(p) = \{q \in D \setminus \{p\} \mid d(p,q) \leq k\text{-}distance(p)\} \quad (1)$$

Note:

D is all the objects, and p is the object we are calculating its k neighborhood. According to the definition, p is not included in its k neighborhood.

Definition 3 (Reachability distance between two objects)**.** *Reachability from one object to another is the maximum distance between two objects or the distance of the k-th neighbor of the second object (o):*

$$reach\text{-}dist_k(p,o) = max\{k\text{-}distance(o), d(p,o)\} \quad (2)$$

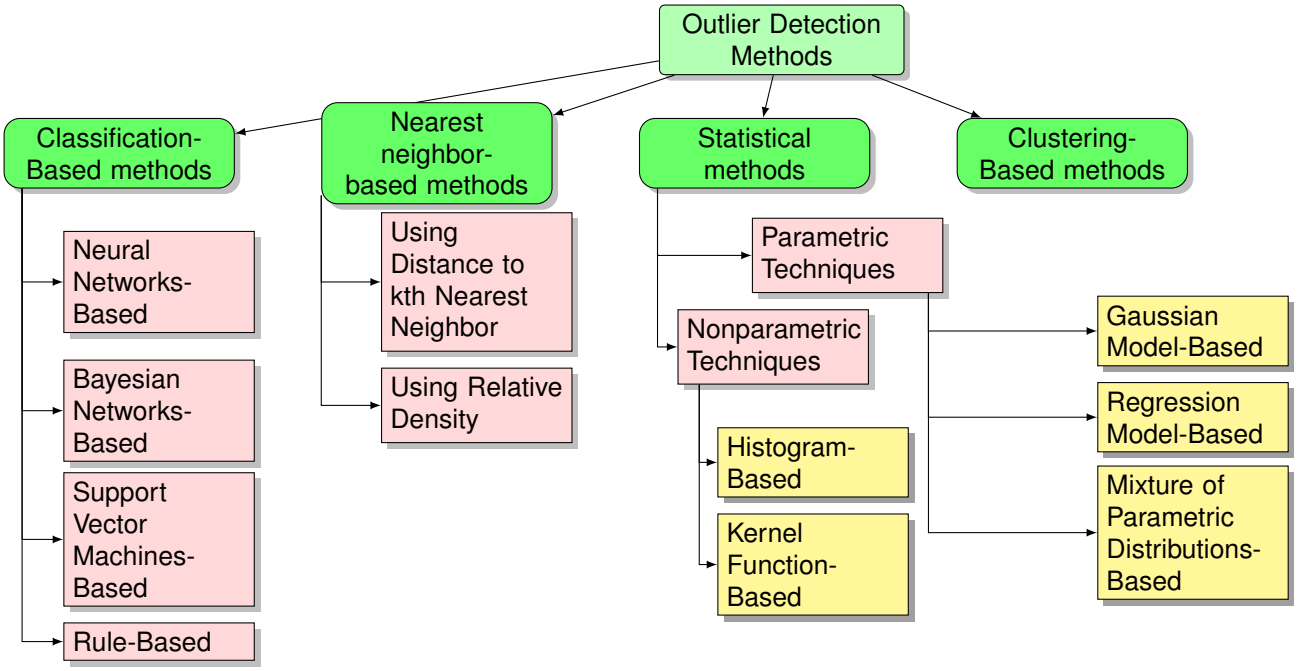


Fig. 1. Four subcategories for outlier detection methods

Definition 4 (Local Reachability Density (lrd) of an object)**.** *The local reachability of p is defined as follows:*

$$lrd_{MinPts}(p) = \frac{\mid N_{MinPts}(p) \mid}{\sum_{o \in N_{MinPts}(p)} reach\text{-}dist_{MinPts}(p,o)} \quad (3)$$

Note:

$NMinPts(p)$ is the k neighborhood of p. Therefore, Formula 3 investigates the correlation of an object locally. The larger this value is, the object and its neighbors are denser.

Definition 5 (Local Outlier Factor (LOF))**.** *LOF is defined using Formula 4.*

$$LOF_{MinPts}(p) = \frac{\sum_{o \in N_{MinPts}(p)} \frac{lrd_{MinPts}(o)}{lrd_{MinPts}(p)}}{\mid N_{MinPts}(p) \mid} \quad (4)$$

As seen, the result is a relative number showing the outlierness. The greater this value is, the more outlier the object is.

B. Shortcoming of the LOF in dense clusters

The shortcoming of the LOF is that it overemphasizes outlier detection around significantly dense clusters. This algorithm, due to its relative analysis of outlierness in a dataset, always encounters difficulties in significantly dense clusters. In fact, it detects too many objects as outliers around dense clusters. As seen in Figure 2 , the LOF algorithm has detected a lot of outliers around $C1$ and $C2$ clusters; that is because in dense clusters, due to the small distance between objects, the denominator of the fraction in Formula 3 approaches zero, and the outlierness score of these objects is large; therefore, most objects distant from those around dense clusters are selected as outliers, and other objects are less likely to be selected as outliers.

Algorithm 1 LOF pseudo code

Input: The number of neighbors (k) and a dataset
Output: An array ordered according to the outlierness scores of each object

Step 1: measuring the distance between all objects
Step 2: Calculating k neighborhood for each object
Step 3: Calculating reach -distk(p, o) for all objects
Step 4: Calculating lrdMinPts(p) for all objects
Step 5: Calculating LOFMinPts(p) for all objects
Step 6: Ordering the values of LOF for all objects and selecting an ideal number of objects with the highest outlierness scores

C. The proposed solution to improve the LOF algorithm

As seen in Figure 3, in the proposed method, to overcome the problem in the LOF algorithm around dense clusters, we generate a random point far from the objects of the dataset, and then, this object is added to the neighbors of all objects. To determine the random object, first, the mean of the values and size of each dimension (maximum and minimum of each dimension) is calculated, and the value of the random object is randomly generated within the range of 2-20 times as much as the size of each dimension farther than the mean of that dimension. This innovative idea adds a great

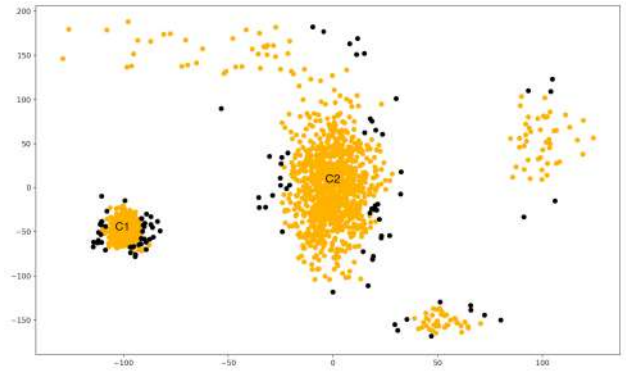


Fig. 2. Performance of LOF in dense clusters (LOF algorithm is concentrated around $C1$)

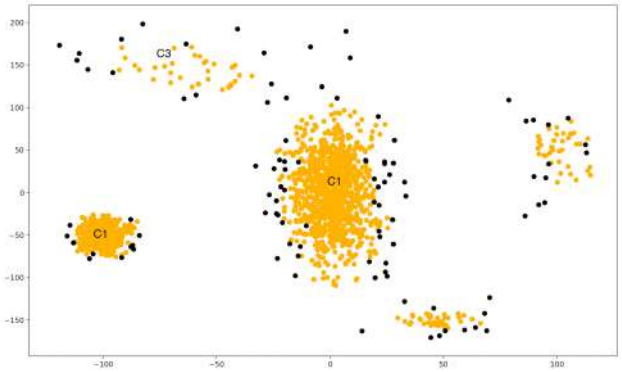


Fig. 4. Performance of the proposed algorithm in dense clusters (The proposed algorithm also covers the objects in $C3$)

value to the distance between each object and its neighbors. Therefore, large and close values are added to the numerator and denominator of Formula 4 since the distances between two neighboring objects and the random object are two large and close values one of which placed in the numerator and the other placed in the denominator. As a result the issue of small distances in dense clusters is solved in Formula 4. The improved LOF algorithm only changes Formula 1, and the new formula is restated as Formula 5; the rest of the formulas are the same as those in the LOF algorithm.

$$N_{k\text{-}distance(p)}(p) = \{q \in D \setminus \{p\} \mid d(p,q) \leq k\text{-}distance(p)\} \cup \{RandomPoint\} \quad (5)$$

Note:

The proposed algorithm is not sensitive to the random point since after various runs, (10-20 runs) with different random points, the detection rate of the algorithm does not significantly change.

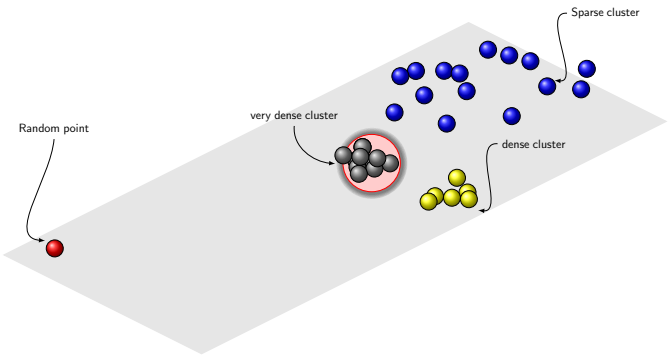


Fig. 3. Adding a random point far from the dataset to the neighborhood of all objects

As seen in Figure 4, the improved LOF algorithm, in addition to detecting outliers around $C1$ and $C2$ clusters, has detected outliers around sparser clusters such as $C3$. This indicates the efficiency of the LOF algorithm.

IV. The Results of the Proposed Algorithm

In this section, we introduce some real datasets used in this paper first, then evaluate our proposed algorithm with LOF algorithm according to precision and AUC criteria.

A. Datasets

To assess the proposed algorithm, first we introduce some datasets. The datasets based on which we evaluated the proposed algorithm include real datasets [14]. Each dataset has features such as the number of objects and the percentage of outliers, which are presented in Table I.

TABLE I
CHARACTERISTICS OF THE REAL DATASETS

dataset	Samples	Dimension	outliers (present)
Annthyroid	7200	6	534(7.42%)
Arrhythmia	452	274	66(15 %)
BreastW	683	9	239(35 %)
Cardio	1831	21	176(9.6 %)
Letter	1600	32	100(6.25%)
Glass	214	9	9(4.2%)
Ionosphere	351	33	126(36 %)
Lympho	148	18	6(4.1 %)
Mnist	7603	100	700(9.2 %)
Musk	3062	166	97(3.2 %)
Optdigits	5216	64	150(3 %)
Pima	768	8	268(35%)
Satellite	6435	36	2036(32%)
Satimage	5803	36	71(1.2 %)
Thyroid	3772	6	93(2.5%)
Vowels	1456	12	50(3.4%)
WBC	278	30	21(5.6%)
Wine	129	13	10(7.7%)

B. Results based on AUC and Precision criteria

The proposed algorithm and the LOF algorithm were evaluated based on some real datasets. Two criteria of Precision and AUC [15] were used to compaire the proposed algorithm with LOF algorithm. The results of this evaluation are presented in Tables II and III. The results indecate that the proposed algorithm has better precision and AUC in many real datasets, and it has less efficiency in a few datasets. The improvements in outlier detection rate indicate that the issue of dense clusters exists in many datasets.

V. Conclusion

In the present study, first the topic of outliers and their vast applications were introduced. Then, outlier detection methods were reviewed. Subsequently, one of the most popular outlier detection algorithms, LOF, was explained in detail. Afterward, one of the shortcomings of this algorithm in dense clusters were stated, and the innovative idea of using random point was presented to overcome this shortcoming. This led to the introduction of the new proposed algorithm. Next, after introducing some real datasets, this new method was compared with the LOF algorithm, showing the competitiveness of the new algorithm and a relative improvement compared with the LOF algorithm.

TABLE II
Comparison of the improved LOF algorithm with LOF according to Precision

dataset	LOF Algorithm	Improved LOF
Annthyroid	30.33 %	29.38 %
Arrhythmia	45.83 %	48.33 %
Breastw	24.60 %	27.00 %
Cardio	18.66 %	23.07 %
Letter	39.65 %	39.25 %
Glass	15.00 %	15.55 %
Ionosphere	74.40 %	76.35 %
Lympho	56.66 %	58.33 %
Mnist	30.62 %	38.14 %
Musk	1.75 %	2.01 %
Optdigits	8.13 %	7.90 %
Pima	44.51 %	46.21 %
Satellite	39.10 %	40.51 %
Satimage	7.18 %	18.02 %
Thyroid	89.51 %	89.77 %
Vowels	35.20 %	35.90 %
WBC	47.85 %	49.76%
Wine	82.00 %	86.00 %
Average	38.38 %	40.64 %

TABLE III
Comparison of the improved LOF algorithm with LOF according to AUC

dataset	LOF Algorithm	Improved LOF
Annthyroid	72.01 %	70.44 %
Arrhythmia	79.33 %	80.24 %
Breastw	45.65 %	46.34 %
Cardio	68.50 %	72.16 %
Letter	84.63 %	84.20 %
Glass	76.10 %	77.00 %
Ionosphere	89.12 %	90.35 %
Lympho	96.36 %	96.10 %
Mnist	72.13 %	73.06 %
Musk	37.14 %	38.89 %
Optdigits	50.21%	49.77 %
Pima	59.91 %	61.24 %
Satellite	55.68 %	56.79 %
Satimage	66.11 %	72.12 %
Thyroid	89.51 %	89.77 %
Vowels	93.51%	93.49 %
WBC	93.12 %	93.58 %
Wine	96.88 %	97.86 %
Average	73.66 %	74.63 %

A. Suggestions for further studies

- The proposed algorithm is not sensitive to random point. However, optimizing random objects to achieve the optimal efficiency of the proposed algorithm is recommended.
- After several analyses of datasets, it seems that using several random points in the proposed algorithm leads to the improvement of its efficiency. Therefore, this idea can be investigated in future studies.

REFERENCES

[1] F. E. Grubbs, "Procedures for detecting outlying observations in Samples," vol. 11, no. 1, 1969, pp. 1–21.

[2] K. Ord, "Outliers in statistical data : , 1994, 3rd edition, (John Wiley & Sons, Chichester), 584 pp., [UK pound]55.00, ISBN 0-471-93094-6," *International Journal of Forecasting*, vol. 12, no. 1, pp. 175–176, 1996.

[3] S. Ghosh and D. L. Reilly, "Credit card fraud detection with a neural-network," in *Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences*, vol. 3. IEEE, 1994, pp. 621–630.

[4] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly Detection: A Survey," *ACM Comput. Surv.*, vol. 41, no. 3, pp. 15:1–15:58, 2009.

[5] T. Fawcett and F. Provost, "Activity monitoring: Noticing interesting changes in behavior," in *Proceedings of the fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 1999, pp. 53–62.

[6] J. Lin, E. Keogh, A. Fu, and H. Van Herle, "Approximations to magic: Finding unusual medical time series," in *Computer-Based Medical Systems. 18th IEEE Symposium on Proceedings.* IEEE, 2005, pp. 329–334.

[7] E. Keogh, J. Lin, S.-H. Lee, and H. Van Herle, "Finding the most unusual time series subsequence: algorithms and applications," *Knowledge and Information Systems*, vol. 11, no. 1, pp. 1–27, 2007.

[8] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, "LOF: identifying density-based local outliers," in *ACM Sigmod Record*, vol. 29, no. 2. ACM, 2000, pp. 93–104.

[9] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*. John Wiley & Sons, 2012.

[10] F. J. Anscombe, "Rejection of outliers," *Technometrics*, vol. 2, no. 2, pp. 123–146, 1960.

[11] E. Eskin, "Anomaly detection over noisy data using learned probability distributions," in *In Proceedings of the International Conference on Machine Learning*. Citeseer, 2000.

[12] G. Gan and M. K.-P. Ng, "k-means clustering with outlier removal," *Pattern Recognition Letters*, vol. 90, pp. 8–14, 2017.

[13] M. Ester, H. P. Kriegel, J. Sander, and X. Xu, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise," *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining*, pp. 226–231, 1996.

[14] S. Rayana, "ODDS Library," 2016. [Online]. Available: http://odds.cs.stonybrook.edu

[15] A. Zimek, E. Schubert, and H.-P. Kriegel, "A survey on unsupervised outlier detection in high-dimensional numerical data," *Statistical Analysis and Data Mining*, vol. 5, no. 5, pp. 363–387, 2012.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Improving Image Captioning with Local Attention Mechanism

Zahra Famil Stattari[1], Hassan Khotanlou[2], Elham Alighardash[3]

[1]*RIV Lab., Department of Computer Engineering, Bu-Ali Sina University, Hamedan, Iran,* z.familsattari@eng.basu.ac.ir
[2]*RIV Lab., Department of Computer Engineering, Bu-Ali Sina University, Hamedan, Iran,* khotanlou@eng.basu.ac.ir
[3]*RIV Lab., Department of Computer Engineering, Bu-Ali Sina University, Hamedan, Iran,* e.alighardash@eng.basu.ac.ir

***Abstract*— Image caption generation is an interdisciplinary field of research in machine vision and natural language processing. Based on the results of studies and reported accuracy, this is a difficult task for the machine to perceive the image like a human. Most of the proposed methods in the field of image description production follow the encoder-decoder framework. In these methods, each word is generated based on the characteristics of the image and the previously generated words. Recently, the attention mechanism, which usually creates a spatial map that highlights the image areas associated with each word, is widely used in researches. In this paper, the encoder-decoder framework is used. The encoder part of the model will use ResNet to extract the global features and the decoder part consists of three important parts: Attention-LSTM, Language-LSTM, and Attention-layer. The attention mechanism uses local evidence to better demonstrate features and reasoning in the generation of image descriptions. Our method was able to improve the evaluation metrics of METEOR, ROUGH well.**

***Index Terms* —Deep learning; image captioning; attention mechanism; encoder-decoder**

I. INTRODUCTION

The purpose of automatic caption generation is to produce content-rich descriptions that are very similar to human descriptions. Image captioning is a vital vision task that has many applications in various fields of computer science such as video retrieval. Moreover, it has asked for use in other domains like digital libraries, education, commerce, and biomedicine. Image captioning can be described as an interdisciplinary research domain that has assimilated computer vision and natural language processing (NLP) [1]. It faces more challenges in comparison with image classification, object recognition, and many other computer vision tasks. This event is due to the need for general and immense perception of the entities within an image and the relationships between them.

The decoder-encoder model is typically used to image captioning. The pixel-level information is fed to a CNN (Convolution Neural Network) decoder to encode it by densing. Then, the decoder is applied to translate information of output of CNN into natural languages. This type of architecture has achieved significant improvements over previous architectures [2]. In the mentioned method, specifically, each element of model output is generated by applying the visual features and previous words that were created in the sequence. Recently, the attention mechanism has been examined in the CNN features map, which typically creates a spatial map that highlights the image regions associated with each generated word [3, 4].

Most methods of image captioning try to improve the evaluation metrics BLEU, BLEU is the geometric mean of precision scores of the obtained N-grams multiplied by brevity penalty, but in this method, we try to improve the criteria of METEOR. METEOR is the tunable mean of two other parameters (precision and recall) given between matching unigrams in a sentence. In addition, it uses synonyms and paraphrase matching. METEOR fixes several BLEU deficiencies, such as recall evaluating and not explicitly word matching. N -gram-based metrics work well when there is considerable overlap between reference and candidate sentences. n-gram based measures work reasonably well when there is a significant overlap between reference and candidate sentences; However, when common words are scarce, they do not recognize the semantic similarity [5].

The contributions of this study can be summarized as follows:

- A novel architecture is designed that consists of three main parts:
    1) The image feature extractor applies a pre-trained ResNet architecture on the ImageNet database.
    2) The LSTM-Attention layer, which includes Local Attention that is derived from the spatial attention mechanism.
    3) The Language-LSTM layer that generates words.
- The aim is to improve the results of analyzing the Flickr8k dataset and is considered based on the ROUGE and METEOR evaluation metrics.

II. RELEATED WORK

A. Encoder-Decoder based structure

The basis of most in-depth methods is the encoder-decoder framework. The simple neural network-based models as the elementary end-to-end models were not able to generate considerable results for image captioning. So, emerging of encoder-decoder architectures, raised from the neural machine translation approaches, changed the condition. In these networks, the output of captioning is a sequence of words that are derived from a two-level framework including a CNN as a global image features extractor and an LSTM that accepts the output of CNN to generate the final caption.

The Neural Image Caption Generator (Google NIC) was a sample of encoder-decoder models that were proposed by Vinyals et al. [6]. In this framework, GoogleNet plays the role of the encoder and extracts visual features of an input image, and then the intermediate result is passed to an LSTM to generate the ultimate sentence. Jia et al. [7] proposed a novel LSTM model. Using the global semantic information of the image as additional information to LSTM cells. This makes the generated captions have higher quality. Fang et al. [8] proposed a method for generating image captions. This method extracts features of test images based on word detectors from training image captions. The related features were the input of a maximum entropy language model to generate a sentence, and then a re-ranking has done for the final output. The study of Lebret et al. [9] relied on the Markov model for generating captions by using the embedded expressions. This method was a bilinear model that tried to learn the common space of image features and textual and syntactic expressions such as prepositional phrases, verb phrases, and noun phrases.

B. Attention based structure

The attention mechanism can focus on different elements of sequence or regions of an image dynamically. Many studies mentioned that attentive methods emphasized on more fine-grained visual features to generate captions by relating salient image parts and words. Xu et al. [4] introduced an attention-based image captioning by a multi-attention approach that merged hard attention with a soft attention mechanism. This method can automatically pay attention to the highlighted contents of an image during sentence generation. The proposed model of Pedersoli et al. [10] applied region-based attention mechanisms. In this method, the relations between the generated words and the image regions were modeled. Lu et al. [3] developed a model of adaptive attention with the visual sentinel, which decides whether or not to pay attention to the image at any time. Anderson et al. [11] offered the idea of combining bottom-up and top-down attention mechanisms to generate caption: The bottom-up selected some image regions that are integrated by convolution property vectors. Then, top-down attention ranked these regions by weighting them during generation. Deng et al. [12] proposed a method in 2020 that uses the DenseNet model [13] to extract image properties. At the same time, the sentinel gate is adjusted by the attention mechanism to decide which image-specific information to use to generate the word. Li et al. [14] Proposed a local-global attention mechanism to include local features extracted from detected objects by a pre-trained object detector. S.Cao et al.[15] proposed the IGGAN model, for unsupervised training that combines the display of multi-scale properties and object relationships. Chen et al. [16] proposed a method to combine spatial attention and channel-wise attention mechanism. Marcella et al. [17] proposed the M^2 transformer architecture, which improved image encoding and language generation. The high computational cost and complexity of training are the weaknesses of this method. Similarly, Hey et al. [18] modified the encoding and decoding transformer for image captioning, and also optimized the loss function.

C. Other Work

Zhang et al. [19] suggested the use of reinforcement learning (RL) in the training phase for direct optimization of evaluation metrics. Auli et al.[20] proposed a new REINFORCE model by adding an unrelated baseline. Recently, the use of generative model–generative adversarial nets (GANs) [21] using the attention mechanism has become common. Chen et al. [1] proposed a new discriminator model to filter human-generated from machine-generated captions. Focusing on high-level semantic features in caption generation, such as concepts, and structural semantics is the property of enhanced semantic methods. Wu et al. [22] study is a sample of applying high-level semantics in the encoder-decoder framework. They confirmed the effectiveness and positive impact of explicit representation of features for image description.

III. PROPOSED METHOD

As explained in the introduction, the most common models of image captioning are based on the encoder-decoder framework. Our model also follows encoder-decoder architecture. Typically, the encoder-decoder architecture uses a CNN as an encoder and an RNN as a decoder. The general architecture is shown in Fig. 1, which uses a CNN (ResNet101) module to extract features, called an encoder. The next module includes a decoder cell that contains the Attention-LSTM, Attention-Layer, and the Language-LSTM module.

As seen in Fig. 1, image I is given as input, and the image captioning model should generate a semantically and syntactically correct sentence, $S = \{y_1, y_2, \ldots, y_L\}$, where i-th word of the sentence of S is y_i and L shows caption length, as follows.

$$F: I \rightarrow S \tag{1}$$

When the model is trained by maximum-likelihood estimation (MLE), the ground-truth textual description of the image up to time step t−1, $\{ y_1^*, \ldots, y_{t-1}^* \}$ will be given to the model, and it is trained to maximize the likelihood of y_t^* as the ground-truth element of time step t. In other word LSTM generates the word probability distribution in time t using the pre-existing sequence of words of ground-truth up to time t -1 and visual features. This model applies cross-entropy loss for optimization [23].

$$L_{CE}(\theta) = -\sum_{t=1}^{T} log\left(p_\theta (y_t^* | v, y_1^*, \ldots, y_{t-1}^*) \right) \quad (2)$$

Theta displays the weights that can be learned in the image captioning model.

The θ parameter shows the learnable weights of the model. The model generates a distribution of every word conditioned on the pre-existing sequence that was generated before and visual features because the ground-truth captions are not available in the form of $p_\theta(y_t | v, y_1, \ldots, y_{t-1})$.

According to Fig. 1, to generate a caption for the image that is syntactically and semantically correct, we follow the following steps.

A. Image Encoder

This architecture uses ResNet101 to extract and detect image spatial features. It is a pre-trained model already trained on the ImageNet dataset. Image global features, $V = \{ v_1, v_2, \ldots, v_k \}$, covers more detail compared to the input image.

$$V = CNN_{CONV}(I) \quad (3)$$

where I is the input image, CNN_{CONV} is the last convolutional layer of ResNet101, vi is a n-dimensional vector that represents features of an input image region, and k is the number of this region.

B. Decoder-Cell

Our decoder cell consists of three main modules: Attention-LSTM, Attention-Layer, and Language-LSTM. In this way, better image understanding has been targeted.

Attention-LSTM: This module mainly attends to the language-specific context and pre-generated textual information, and then continues to produce the distribution of attention on all image areas [24]. As shown in Figure 1. it means after concatenation of the image global feature, $v_g = \{ v_1. v_2. \ldots. v_k \}$, with the vector of previously generated caption ($W_e y_{t-1}$), the result feeds to Attention-LSTM, as follows:

$$h_t^1 = LSTM(LSTM([v_g . W_e y_{t-1}])) \quad (4)$$

Attention-Layer: A layer of attention is proposed that uses local evidence to more effectively demonstrate features and reasoning in image captions.

The overall architecture of the proposed Attention-Layer model is shown in Fig. 2. In this study, Attention is a dense layer that manipulates the lower-level features of a local region to create a spatial map that highlights image regions based on LSTM's feedback information. First, the extracted image features (v_g) are transferred into a new feature space (f), then the outputs of Attention-LSTM(h_t^1) are embedded into a new feature space(g) to calculate local attention.

$$f(v) = W_f v \quad (5)$$

$$g(h_t^1) = W_g h_t^1 \quad (6)$$

$$a_{i.t} = \tanh(f+g) \quad (7)$$

$$\alpha_t = \text{softmax}(a_t) \quad (8)$$

Finally, the result of the local attention module is calculated by:

$$c_t = v_g \times \alpha_t \quad (9)$$

Language-LSTM: The input of this module consists of selected visual features (c_t) and the embedding of previously generated caption $W_e y_{t-1}$, them feed to Language-LSTM, as follow:

$$h_t^2 = LSTM([c_t, W_e y_{t-1}]) \quad (10)$$

Finally, the sum of Attention-LSTM and Language-LSTM is fed to a softmax function to obtain the word probability distribution:

$$p(y_t / y_{1:t-1}) = softmax(h_t^1 + h_t^2) \quad (11)$$

IV. Evaluation

The experimental evaluation results of the proposed method on the Flickr8k [25] image captioning dataset are discussed in this section, and also they are compared with several novel advanced methods.

A. Dataset

Flickr8k contains 8,000 data, including 6,000 training data, 1,000s test data, and 1,000 verification data.

B. Evaluation metric

The evaluation metrics in the image captioning are the same as the evaluation metrics in machine translation.

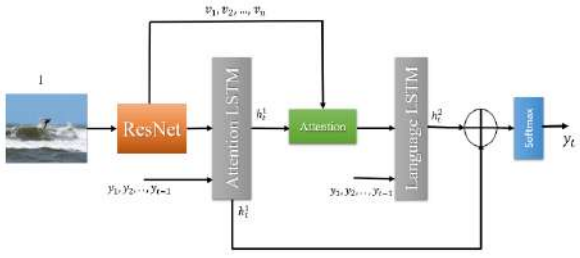


Figure 1. The overall architecture proposed model, it includes (ResNet, Attention-LSTM , Attention-Layer, Language-LSTM)

The evaluation metrics we use includes: BLEU-1, BLEU-2, BLEU-3, BLEU-4 [26], METEOR [27] and ROUGE-L [28]. BLEU (1-4) is one of the first evaluation metrics, BLEU is the geometric mean of precision scores of the obtained N-grams multiplied by brevity penalty.

The METEOR evaluation metric is the harmonic mean of precision and recall of uni-gram. This metric contains several features that other metrics do not have, such as root matching and synonyms. The ROUGE evaluation metric was calculated by comparing the overlap of n-grammars, word sequences, and word pairs. In this paper, we use ROUGE -L, which measures the longest common sequence between a pair of sentences.

C. *Implementation Details*

In the proposed model, we use the ResNet-101 model pre-trained on ImageNet for extracted image features. The size of the extracted features is (2048). The number of LSTM-hidden states are 128, word embedding dimensions are 128 and the number of kernel dense layers are 2048. We used the ADAM optimizer with a learning rate of 0.001. We have trained the model in 70 epoch and batch size 5.

D. *Comparisons*

In this work, we tried to improve the evaluation metrics of METEOR, ROUGE, as you can see in Table 1, Table 2. we improved METEOR, ROUGE because the METEOR metric fixes several BLEU flaws and uses synonyms to match sentences, which is why they are so important. First, we compare our proposed model with the base model (IGCA) on dataset Flickr8k. The results of model in Table. 1 show that the METEOR and ROUGE metrics have been improved.

In the next section, we will compare the proposed method with the novel methods. And as you can see in Table. 2, the METEOR and ROUGH metrics have improved.

E. *Qualitative analysis*

To intuitively show the quality of the proposed model, we show the loss diagram in the training data and the verification data in Fig. 3.

We also show an example of the generated descriptions in Figure 4, as you can see the generated descriptions are of good quality and have described the image well.

V. COCLUSION

As explained, to generate a high-quality image description we have introduced a new model of encoder-decoder. In this model, the ResNet network was used to extract the global features of the image, and also the attention mechanism was used, which makes the features more efficient. Our method was able to improve the evaluation metrics of METEOR, ROUGH well, which have scores of 24.4 and 51.6, respectively, in the Flickr8k database.

Our future work will examine the addition of object recognition and global attention mechanisms to the model. Adding object recognition may present challenges, including the possibility of creating a bounding box that covers all areas of the image.

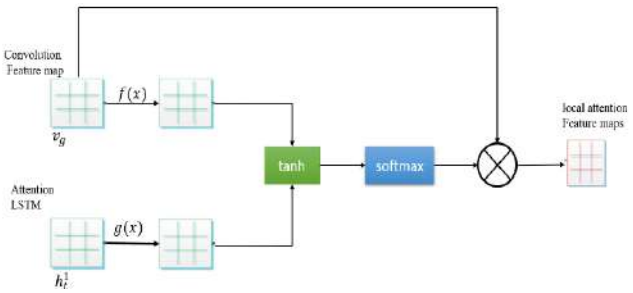


Figure 2. The proposed local attention mechanism

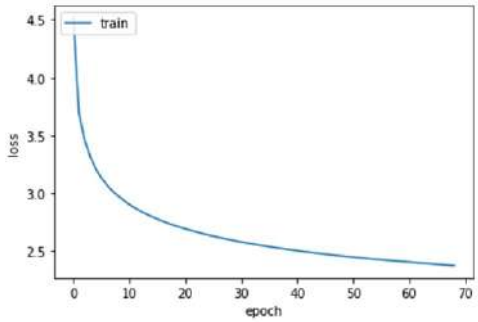


Figure 3. Train loss of our model

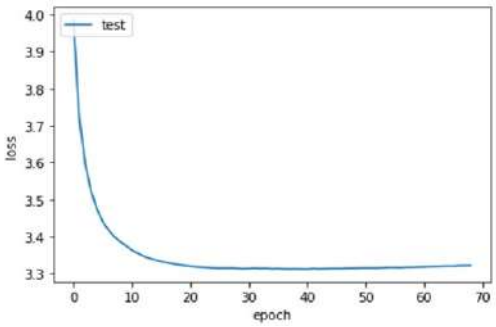


Figure 4. Test loss of our model

TABLE I. COMPARE PORPOSED MODEL WITH BASE MODEL

Method	Flickr8k					
	B-1	***B-2***	***B-3***	***B-4***	***METEOR***	***ROUGE-L***
IGCA[24]	68.9	47.4	33.4	23.8	21.4	48.8
Our	60.4	41.3	27.7	18.7	**24.4**	**51.6**

TABLE II. COMPARE PORPOSED MODEL WITH OTHER MODEL

Method	Flickr8k					
	B-1	***B-2***	***B-3***	***B-4***	***METEOR***	***ROUGE-L***
Xu et al.[4]	67	45.7	31.4	21.3	20.3	-
Fu et al.[29]	63.9	45.9	31.9	21.7	20.4	47
Li et al.[14]	57.2	37.9	23.9	14.8	16.6	41.9
Zhu et al.[30]	63.2	44.8	31.3	21.5	20.4	-
Wang et a.[23]	68.9	57.4	33.4	23.8	21.4	48.8
Our	60.4	41.3	27.7	18.7	**24.4**	**51.6**

Two dogs are running in a field of grass

A black dog is running through the snow

A basketball player in a white uniform is playing with a ball in the background

Figure 5. Generated caption with porposed model

REFERENCES

[1] C. Chen, S. Mu, W. Xiao, Z. Ye, L. Wu, and Q. Ju, "Improving image captioning with conditional generative adversarial nets," in Proceedings of the AAAI Conference on Artificial Intelligence, 2019, vol. 33, no. 01, pp. 8142-8150.

[2] A. Karpathy and L. Fei-Fei, "Deep visual-semantic alignments for generating image descriptions," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 3128-3137.

[3] J. Lu, C. Xiong, D. Parikh, and R. Socher, "Knowing when to look: Adaptive attention via a visual sentinel for image captioning," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 375-383.

[4] K. Xu et al., "Show, attend and tell: Neural image caption generation with visual attention," in International conference on machine learning, 2015: PMLR, pp. 2048-2057.

[5] M. Kilickaya, A. Erdem, N. Ikizler-Cinbis, and E. Erdem, "Re-evaluating automatic metrics for image captioning," arXiv preprint arXiv:1612.07600, 2016.

[6] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan, "Show and tell: A neural image caption generator," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 3156-3164.

[7] X. Jia, E. Gavves, B. Fernando, and T. Tuytelaars, "Guiding the long-short term memory model for image caption generation," in Proceedings of the IEEE international conference on computer vision, 2015, pp. 2407-2415.

[8] H. Fang et al., "From captions to visual concepts and back," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 1473-1482.

[9] R. Lebret, P. Pinheiro, and R. Collobert, "Phrase-based image captioning," in International Conference on Machine Learning, 2015: PMLR, pp. 2085-2094.

[10] M. Pedersoli, T. Lucas, C. Schmid, and J. Verbeek, "Areas of attention for image captioning," in Proceedings of the IEEE international conference on computer vision, 2017, pp. 1242-1250.

[11] P. Anderson et al., "Bottom-up and top-down attention for image captioning and visual question answering," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 6077-6086.

[12] Z. Deng, Z. Jiang, R. Lan, W. Huang, and X. Luo, "Image captioning using DenseNet network and adaptive attention," Signal Processing: Image Communication, vol. 85, p. 115836, 2020.

[13] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 4700-4708.

[14] L. Li, S. Tang, Y. Zhang, L. Deng, and Q. Tian, "GLA: Global–local attention for image description," IEEE Transactions on Multimedia, vol. 20, no. 3, pp. 726-737, 2017.

[15] S. Cao, G. An, Z. Zheng, and Q. Ruan, "Interactions guided generative adversarial network for unsupervised image captioning," Neurocomputing, vol. 417, pp. 419-431, 2020.

[16] L. Chen et al., "Sca-cnn: Spatial and channel-wise attention in convolutional networks for image captioning," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 5659-5667.

[17] M. Cornia, M. Stefanini, L. Baraldi, and R. Cucchiara, "Meshed-memory transformer for image captioning," in Proceedings of the IEEE/CVF IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 10578-10587. Conference on Computer Vision and Pattern Recognition, 2020, pp. 10578-10587.

[18] S. He, W. Liao, H. R. Tavakoli, M. Yang, B. Rosenhahn, and N. Pugeault, "Image captioning through image transformer," in Proceedings of the Asian Conference on Computer Vision, 2020.

[19] L. Zhang et al., "Actor-critic sequence training for image captioning," arXiv preprint arXiv:1706.09601, 2017.

[20] M. A. Ranzato, S. Chopra, M. Auli, and W. Zaremba, "Sequence level training with recurrent neural networks," arXiv preprint arXiv:1511.06732, 2015.

[21] I. Goodfellow et al., "Generative adversarial nets," Advances in neural information processing systems, vol. 27, 2014.

[22] Q. Wu, C. Shen, P. Wang, A. Dick, and A. Van Den Hengel, "Image captioning and visual question answering based on attributes and external knowledge," IEEE transactions on pattern analysis and machine intelligence, vol. 40, no. 6, pp. 1367-1381, 2017.

[23] X. Liu, H. Li, J. Shao, D. Chen, and X. Wang, "Show, tell and discriminate: Image captioning by self-retrieval with partially labeled data," in Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 338-354.

[24] C. Wang and X. Gu, "Image captioning with adaptive incremental global context attention," Applied Intelligence, pp. 1-23, 2021.

[25] T.-Y. Lin et al., "Microsoft coco: Common objects in context," in European conference on computer vision, 2014: Springer, pp. 740-755.

[26] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, "Bleu: a method for automatic evaluation of machine translation," in Proceedings of the 40th annual meeting of the Association for Computational Linguistics, 2002, pp. 311-318.

[27] S. Banerjee and A. Lavie, "METEOR: An automatic metric for MT evaluation with improved correlation with human judgments," in Proceedings of the acl workshop on intrinsic and extrinsic evaluation measures for machine translation and/or summarization, 2005, pp. 65-72.

[28] C.-Y. Lin, "Rouge: A package for automatic evaluation of summaries," in Text summarization branches out, 2004, pp. 74-81.

[29] K. Fu, J. Jin, R. Cui, F. Sha, and C. Zhang, "Aligning where to see and what to tell: Image captioning with region-based attention and scene-specific contexts," IEEE transactions on pattern analysis and machine intelligence, vol. 39, no. 12, pp. 2321-2334, 2016.

[30] X. Zhu, L. Li, J. Liu, Z. Li, H. Peng, and X. Niu, "Image captioning with triple-attention and stack parallel LSTM," Neurocomputing, vol. 319, pp. 55-65, 2018.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Improving The Obtained Results Of Monte Carlo Simulation For Interval Linear Programming Problem By Using Particle Swarm Optimization

1st Moslem Javanmard, 2nd Hassan Mishmast Nehi

1st Math teacher with a PhD, *Education Department*, Zabol, Iran, javanmard.moslem@gmail.com
2nd Mathematics Department, *University of Sistan and Baluchestan*, Zahedan, Iran, hmnehi@hamoon.ac.ir

***Abstract*— In this paper, we consider interval linear programming problem with equality constraints. Since each interval linear programming problems have an infinite number of characteristic models. We are just looking for the set of all optimal values (best and worst value of the objective function). Computing the best value of the objective function is easy, but obtaining the worst value of the objective function is NP-hard. In this research, firstly, we determine range of optimal values of the objective function and then we obtain it, via particle swarm optimization method. Finally, we provide one numerical example to verify and compare the obtained results of it with the obtained results through Monte Carlo simulation and the exact optimal values.**

***Index Terms*— exact optimal values; interval linear programming; Monte Carlo simulation; particle swarm optimization**

I. Introduction

In the recent decade, some methods for solving (type-1 fuzzy, type-2 fuzzy, interval type-1 fuzzy, interval type-2 fuzzy) linear programming problems based on interval linear programming (ILP) models have been proposed, i. e., [2, 12–16]. As well as, many methods have been suggested to solve ILP problems, i. e., [2, 4, 9]. Some methods are based on interval arithmetic and extensions of the simplex algorithm to the case of interval data [6, 7], while another use a direct inspection [10]. Also, some of them are focused on basis stability [4]. Under the assumption of basis stability, it is possible to obtain the optimal solution set of ILP, but still obtaining optimal values of it, are too complicated. Therefore, in this paper, a new approach to solve the ILP model is discussed. Also, to fix these complexities, we use particle swarm optimization (PSO) for solving the ILP problem. A novel evolutionary computational technique, PSO has been proposed by Kennedy, "et al" [8]. The PSO is the generic computational technique espoused from the progression of biological life in the natural humanity [9]. In this paper, we use PSO method for solving the ILP problems with equality constraints. Obtaining the best and worst value of the objective function of the ILP model is considered. Also, we determine the optimal values with help PSO method and then we compare the obtained optimal values through it with the obtained optimal values through Monte Carlo simulation [1] and the exact optimal values.

II. Analysis of the Optimal Values

An interval number $x^{\pm}$ is generally shown with $[x^-, x^+]$ where $x^- \leq x^+$. If $x^- = x^+$, then $x^{\pm}$ will be degenerate. An interval vector is defined as $b^{\pm} = [b^-, b^+] = \{b^- \leq b \leq b^+\}$, where b^- and b^+ are two numbers in $\mathbb{R}$ and $b^- \leq b^+$. We denote the center and the radius of $b^{\pm}$ with $b^c = \frac{1}{2}(b^+ + b^-)$, and $\Delta_{b^{\pm}} = \frac{1}{2}(b^+ - b^-)$, respectively. Also, an interval matrix is defined as $A^{\pm} = [A^-, A^+] = \{A^- \leq A \leq A^+\}$, where A^- and A^+ are two matrices in $\mathbb{R}^{m\times n}$, and $A^- \leq A^+$. We denote the center and the radius of $A^{\pm}$ with $A^c = \frac{1}{2}(A^+ + A^-)$, and $\Delta_{A^{\pm}} = \frac{1}{2}(A^+ - A^-)$, respectively. The set of all $m \times n$ interval matrices is denoted by $\mathbb{IR}^{m\times n}$. A square matrix $A^{\pm}$ is called regular if each $A \in A^{\pm}$ is non-singular. An interval vector as $x^{\pm} = [x^-, x^+] = \{x^- \leq x \leq x^+\}$, where $x^-, x^+ \in \mathbb{R}^n$ is one column interval matrix.

A general form of ILP model with equality constraints can be expressed as follows:

$$\begin{aligned} &\text{Min } Z^{\pm} = c^{\pm}x^{\pm} \\ &\text{s.t.} \\ &\qquad A^{\pm}x^{\pm} = b^{\pm}, \\ &\qquad x^{\pm} \geq 0, \end{aligned} \tag{1}$$

where the right-hand sides of the constraints and the coefficients of both objective function and constraints are interval matrices (i.e., $b^{\pm}, c^{\pm}, A^{\pm}$). Or, equivalently:

$$\begin{aligned} &\text{Min } z^{\pm} = \sum_{j=1}^{n} c_j^{\pm} x_j^{\pm} \\ &\text{s.t.} \\ &\qquad \sum_{j=1}^{n} a_{ij}^{\pm} x_j^{\pm} = b_i^{\pm}, \quad i = 1, 2, \ldots, m, \\ &\qquad x_j^{\pm} \geq 0, \quad j = 1, 2, \cdots, n, \end{aligned} \tag{2}$$

where the coefficients are equal, $c_j^{\pm} = [c_j^-, c_j^+]$, $a_{ij}^{\pm} = [a_{ij}^-, a_{ij}^+]$, $b_i^{\pm} = [b_i^-, b_i^+]$ and $z^{\pm} = [z^-, z^+]$.

The interval $[z_{opt}^-, z_{opt}^+]$ is called the range of the optimal values of the ILP model (2). Computing of the best value z_{opt}^- is easy, but obtaining the worst value z_{opt}^+ is much more involved [3].

The model $\min\{c^T x | Ax = b, x \geq 0\}$, where $c \in [c^-, c^+] \subseteq \mathbb{R}^n, b \in [b^-, b^+] \subseteq \mathbb{R}^m, A \in [A^-, A^+] \subseteq \mathbb{R}^{m \times n}$ is called B-stable with basis B, if B has been an optimal basis for any characteristic model. B-stability (basis stability) conditions are as [5]: 1. A_B is non-singular or regular (regularity), 2. $A_B^{-1}\mathbf{b} \geq 0$, (feasibility), 3. For minimization problem if $\mathbf{c}_N^T - \mathbf{c}_B^T A_B^{-1} A_N \geq 0^T$, then A_B is optimal (optimality). Under the assumption of basis stability, it is possible to obtain the optimal solution set of ILP.

Let Y_m be the set of all $\{-1, 1\}$ m-dimensional vectors, i.e. $Y_m = \{y \in \mathbb{R}^m | |y| = e\}$, where e=$(1, 1, ..., 1)^T$. For a given $y \in \{-1, 1\}^T$, by $T^y = diag(y_1, y_2, ..., y_m)$, we denote the corresponding diagonal matrix. For each $x \in \mathbb{R}^n$, we denote its sign vector sign(x) by,

$$sign(x_i) = \begin{cases} 1 & if \quad x_i \geq 0 \\ -1 & if \quad x_i < 0, \end{cases}$$

where, $i = 1, 2, ..., n$.

Then we have $|z| = T_z x$, where $z = sign(x) \in Y_n$. Let an interval matrix $A^{\pm} = [A^-, A^+] = [A^c - \Delta_{A^{\pm}}, A^c + \Delta_{A^{\pm}}]$ and an interval vector $b^{\pm} = [b^-, b^+] = [b^c - \Delta_{b^{\pm}}, b^c + \Delta_{b^{\pm}}]$. For each $y \in Y^m$ and $z \in Y^n$, A_{yz} and b_y are defined as follows [5]:

$\mathrm{A}_{yz} = A^c - T_y \Delta_{A^{\pm}} T_z, b_y = b^c + b_y \Delta_{b^{\pm}}$.

[3] The best and the worst values of the objective function of the ILP model (1) could be obtained as follows:

$$z_{opt}^- = \{c^{-T} x | A^- x \leq b^+, A^+ x \geq b^-, x \geq 0\},$$

$$z_{opt}^+ = sup_{y \in Y_m} f(A_{ye}, b_y, c^+),$$

where, f(A, b, c)=inf$\{c^T x | Ax = b, x \geq 0\}$.

[3] Computing the worst value of the objective function of the ILP model (1) is NP-hard.

III. Overview of PSO method

Particle swarm optimization (PSO) is a metaheuristic method inspired by swarm behavior Eberhart and Kennedy [8]. A set of particles represents the solution. Particle x is updated based on its own best memory $lbest$ and social best memory $gbest$. The updating rule in PSO also consists of a parameter which controls the ratio of local best and global best. It is named inertia weight, W. In the original PSO, the inertia weight is set as a constant value. However, it is also can be set as a dynamic value. The updating rule is given in following equations:

$V^{new} =$
$W.V^{old} + c_1 rand_1.(lbest - X^{old}) + c_2 rand_2.(gbest - X^{old})$,
$X^{new} = X^{old} + V^{new}$.

Where V is the velocity. Learning rate C_1 controls the influence of $lbest$ while learning rate c_2 controls the influence of $gbest$.

IV. Analysis of validity

Suppose the intervals $z_{opt} = [z_{opt}^-, z_{opt}^+]$ and $z = [z^-, z^+]$ as the exact range and the obtained range by PSO of the optimal values of the ILP model (1), respectively. So, we define the degree of uncertainty for the obtained range as follows:

$$UD(z) = \left(\frac{|z^- - z_{opt}^-| + |z^+ - z_{opt}^+|}{|z_{opt}^+ - z_{opt}^-|} \right) \times 100.$$

This indicator clearly shows that the obtained range is closer to the exact optimal range, if the degree of uncertainty associated with it, is close to be zero percent. As well as, the degree of feasibility for the obtained range is defined as follows:

$$FD(z) = \left(\frac{|(z_{opt}^+ - z_{opt}^-) - (z^+ - z^-)|}{|z_{opt}^+ - z^+| + |z_{opt}^- - z^-|} \right) \times 100.$$

Also, this indicator clearly shows that the obtained range is sufficiently feasible and logical, if the degree of feasibility corresponding with it, is close to one hundred percent. So, the obtained range z is completely valid, if $UD(z) = 0\%$ and $FD(z) = 100\%$.

V. Numerical example and comparison of the results

In this section, we determine the optimal values for the following example with help PSO method and then we compare the obtained optimal values through it with the obtained optimal values through Monte Carlo simulation based on three distribution functions, normal, uniform and beta [1] and the exact optimal values. Moreover, to show validity of PSO method, we use the indicators of the degree of uncertainty and the degree of feasibility.

Consider the interval linear programming

$$\begin{aligned} &\text{Min } \ Z^{\pm} = x_1^{\pm} + x_2^{\pm} \\ &\text{s.t.} \\ &\qquad [2, 3] x_1^{\pm} + x_2^{\pm} = [3, 4] \\ &\qquad [4, 6] x_1^{\pm} + [-3, -2] x_2^{\pm} = [1, 5] \\ &\qquad x_1^{\pm}, x_2^{\pm} \geq 0. \end{aligned} \tag{3}$$

By theorem II, the best and the worst values of the objective function of ILP model (V) are $z^{-}_{opt} = 1$ and $z^{+}_{opt} = 3.1$, respectively. So, the range of exact optimal values of the objective function is $[1, 3.1]$. Also, the obtained optimal values through Monte Carlo simulation based on three distribution functions, normal, uniform and beta for Example (V), are persented in Table (I) [1].

TABLE I
RESULTS OF MONTE CARLO SIMULATION BASED ON THREE DISTRIBUTIONS

Distribution	***Values of z***	$UD(z)$	$FD(z)$
Normal	***[1.0394, 2.8585]***	***13.38 %***	***100%***
Uniform	***[0.9512, 2.9398]***	***9.95%***	***53.30%***
Beta	***[0.9384, 2.8186]***	***16.33%***	***64.08%***

The initial parameters related to PSO method for solving this ILP problem are considered in Table (II). The obtained optimal values for Example (V), through PSO method are presented in Table (III). Also, the values of z resulted from it, are shown in Fig. 1 and Fig. 2.

TABLE II
INITIAL PARAMETERS RELATED TO PSO METHOD

Method	***PSO***
parameter	***Population: 20, Vmax: 10*** ***Inertial weight: 0.2 – 1.2, Iteration: 100***

TABLE III
RESULTS OF PSO METHOD

Method	***Values of z***	$UD(z)$	$FD(z)$
PSO	***[1.009, 3.09]***	***0.90 %***	***100 %***

Fig. 1. The best value of z resulted from PSO method

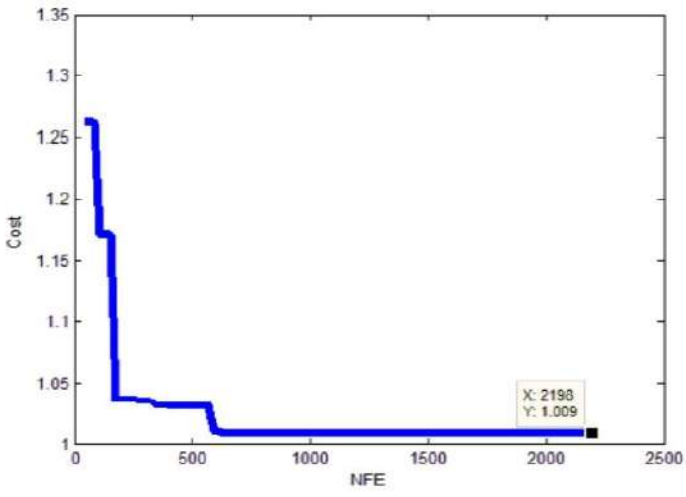
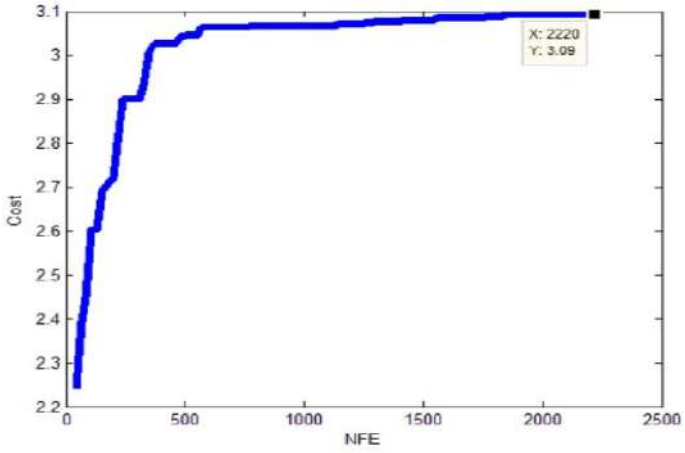


Fig. 2. The best value of z resulted from PSO method



VI. CONCLUSION

Computing the worst value of the objective function is very complicated for ILP models with equality constraints. PSO method has been used to explore the values of the objective function. The obtained results in Example (V), show that the obtained optimal values, through PSO method are much better than from Monte Carlo simulation results. Moreover, the obtained optimal values, through PSO method are much accurate and completely logical, while the obtained optimal values, through Monte Carlo simulation are highly inexact and as well as a large part of them are illogical and impossible. So, it is clear that obtaining optimal range by PSO method is not only possible but also by increasing the number of iterations in its implementation, the exact optimal range can be obtained.

REFERENCES

[1] M. Allahdadi, A. Khaje Golestane, (2015), Monte Carlo Simulation for Computing the Worst Value of the Objective Function in the Interval Linear Programming. Int. J. Appl. Comput. Math. pp 1–10, Springer, India.

[2] M. Allahdadi, H. Mishmast Nehi, H. A. Ashayerinasab, M. Javanmard, (2016), Improving the modified interval linear programming method by new techniques, Information Sciences, 339, pp. 224–236.

[3] M. Fiedler, J. Nedoma, J. Ramik, K. Zimmermann, (2003), Linear Optimization Problems With Inexact Data, New York.

[4] M. Hladik, (2014), How to determine basis stability in interval linear programming, Optim. Lett., 8, 375–389.

[5] M. Hladik, (2012), Complexity of necessary efficiency in interval linear programming and multiobjective linear programming. Optim. Lett. 6, 893–899.

[6] C. Jansson, (1988), A self-validating method for solving linear programming problems with interval input data, In: Kulisch, U., Stetter, H.J. (eds.) Scientific computation with automatic result verification, Computing Suppl. 6, pp. 33–45.

[7] C. Jansson, S. M. Rump, (1991), Rigorous solution of linear programming problems with uncertain data, Oper. Res., 35(2), 87–111.
[8] J. Kennedy, R. C. Eberhart, Y. Shi, (2001), Swarm Intelligence, Morgan Kaufmann, San Francisco, U.S.A.
[9] N. Khilwani, A. Prakash, R. Shankar, M. K. Tiwari, (2008), Fast clonal algorithm, Eng. Appl. Artif. Intell., 21, 106–128.
[10] R. Krawczyk, (1975), Fehlerabschung bei linearer optimierung, In: Interval Mathematics, LNCS 29, pp. 215–222, Berlin.
[11] M. Javanmard, H. Mishmast Nehi, Rankings and operations for interval type-2 fuzzy numbers: a review and some new methods, Journal of Applied Mathematics and Computing, DOI: https://doi.org/10.1007/s12190-018-1193-9, First Online: 31 May 2018.
[12] M. Javanmard, H. Mishmast Nehi, Arithmetic operations and ranking methods for interval type-2 fuzzy numbers; A review and some new methods, 9th International Iranian Operation Research conference, IORC, Shiraz, Iran (2016), 928-933.
[13] M. Javanmard, H. Mishmast Nehi, Solving the interval type-2 fuzzy linear programming problem by the nearest interval approximation, 6th Iranian Joint Congress on Fuzzy and Intelligent Systems, CFIS, Kerman, Iran, DOI: https://doi.org/10.1109/CFIS.2018.8336609, Publisher IEEE, Feb 2018.
[14] M. Javanmard, H. Mishmast Nehi, A Solving Method for Fuzzy Linear Programming Problem with Interval Type-2 Fuzzy Numbers, Int. J. Fuzzy Syst., DOI: https://doi.org/10.1007/s40815-018-0591-3, Published Online: 17 January 2019.
[15] M. Javanmard, H. Mishmast Nehi, Interval type-2 triangular fuzzy numbers; New ranking method and evaluation of some reasonable properties on it, 5th Iranian Joint Congress on Fuzzy and Intelligent Systems, CFIS, **8003587**(2017), 4–6.
[16] M. Javanmard, H. Mishmast Nehi, Solving interval type-2 fuzzy linear programming problem with a new ranking function method, 5th Iranian Joint Congress on Fuzzy and Intelligent Systems, CFIS, **8003586**(2017), 1–3.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Intelligent fault detection of planetary gearbox using vibration signal processing by empirical mode decomposition and an integrated artificial neural network-support vector machine classifier

Mahdi Armoon[1], Marzie Lafouti[2], Amirhossein Nikoofard[3]

[1]*Faculty of Electrical Engineering, K.N.Toosi University of Technology*, mahdi.armoon@email.kntu.ac.ir
[2]*Faculty of Electrical Engineering, K.N.Toosi University of Technology*, lafuti@email.kntu.ac.ir
[3]*Department of Control Engineering, K.N.Toosi University of Technology*, a.nikoofard@kntu.ac.ir

***Abstract*— Nowadays, automatic intelligent maintenance in fault detection in industry, transportation, airspace, and automation industry is one of the most significant issues. Reliability and availability of the systems are important to agricultural and industrial organizations. The main goal is to increase the time of the system without fault and decrease maintenance costs. As human sources are decreased and equipment has become sophisticated, intelligent maintenance should be replaced with old typical methods. Using novel methods in various applications such as using vibration analysis in monitoring, has given wide information about the operation of instruments which leads engineers in programming, control, navigation, and optimization to the best way. In this paper, an effective method based on vibration analysis for fault detection of planetary gearboxes is presented. The signals are taken from the final reducer of Massey Ferguson tractor 285 with healthy and faulty gears. Vibration signals were processed by empirical mode decomposition (EMD). Characteristics of the decomposed signals were provided statistically for the conditions of the gearbox. By using the integrated classification of artificial neural network-support vector machine (ANN-SVM), the classification of different conditions of gearboxes is given. This algorithm can solve the problem of Multi-layer perceptron which is getting into local minima and SVM method turns the problem into a global optimization problem. The accuracy of the proposed classifier for train and test data is reported at 97.14% and 91.11% respectively.**

***Index Terms* — Component, Gearbox; Planetary; Vibration signal; Empirical mode decomposition; Decision integration.**

I. INTRODUCTION

In recent decades, system monitoring is a practical way in order to detect different faults in the industrial instrument which needs high precision in maintenance. These systems are able to detect faults before failure occurs in the instrument or happening injuries to the operator [1]. One of the most commonly used mechanisms used in industrial equipment is gearboxes. Due to the critical role of planetary gearboxes and in the operation of many devices and the many damages that may occur due to their failure, timely determination and detection of faults in the gears are very important. On the other hand, due to the variety of faults and working conditions of most gears that are usually used in conjunction with other moving components, the complex mechanism of engagement of gears, the presence of disturbing factors in the workplace such as noise, the presence of indeterminate components and many other factors, have made fault detection in gears difficult [2]. In addition, the situation in today's industries, along with the characteristics of traditional methods, has made the use of traditional methods such as the use of the five senses, etc., inefficient, and has made the use of modern and automated methods a necessity to monitor the status of industrial equipment.

Various methods in the process of expanding and completing the monitoring of the condition of the gears have been introduced so far, which from the point of view of the physical nature of the received signal can be referred to as vibration analysis [3], sound propagation [4], electric current changes [5] and so on. Vibration analysis is the strongest and most common method for determining the faults of gears of the gearbox systems [6]. In this method, vibrations due to the engagement of

the gears that are transmitted to the body and structure of the gearbox are recorded by a vibration sensor. Vibration signals usually have noise, so to provide a suitable feature vector, different signal processing methods are used to pre-process the vibration signals, among which EMD is one of the most important and efficient methods in fault detection. The empirical mode decomposition method which is used for static signals can provide information in both time and frequency domains simultaneously [7].

A lot of research has been done to diagnose faults in gearboxes based on vibration data and methods based on artificial intelligence, which shows the high efficiency of industrial intelligence in classifying the type of faults in gearboxes [8-11]. It is capable of increasing the accuracy of systems based on industrial intelligence in which several classifiers are combined simultaneously so that the output of one classifier will be used as the input of the next classifier [12,13]. A review of the references showed that the EMD method is strong in extracting faulty information from vibrational signals and also the integration to the extent of making decisions will increase the accuracy of classification. Therefore, this study aims to provide methods for detecting planetary gearbox faults by using vibration signals in the EMD and combining support vector machine (SVM) and artificial neural network (ANN) to get better performance as the result.

In this paper, after reviewing theoretical foundations containing Empirical mode decomposition (EMD), Support vector machine (SVM), and Multi-layer perceptron neural network (MLP) in section 2, the experiment platform, tools for recording data, signal processing, feature extraction, and the integrated model for fault diagnosis are proposed, respectively in section 3. The results of each are given in section 4. Due to having the SVM method in this proposed algorithm's heart, it can solve the problem of getting into local minima which is one of the major disadvantages of the multi-layer perceptron network and the SVM method turns the problem into a global optimization problem. Indeed, this issue is the main reason that this integrated algorithm is chosen to be applied. In addition, the sensitivity of classification depends on SVM's parameters. Finally, in the last section, section 5, the results of these three presented methods have been discussed.

II. A REVIEW OF THEORETICAL FOUNDATIONS

A. Empirical mode decomposition (EMD)

One of the powerful methods of signal processing in the frequency domain is the method of decomposing the signal into intrinsic components. This method, known as empirical mode decomposition or EMD, decomposes a signal into Figure 1, showing the step-by-step method of signal decomposition into a subband and a residue. In this way, first, the up and down signal is taken from the main signal. The signal is then averaged from the top and bottom folds, which is extracted as the subband or IMF, and the rest of the signal is considered as the residual. Again from the residual signal, the top and bottom cover is taken, and this process continues until the top cover and bottom cover match. Eventually, the signal will be decomposed into the original subbands, and a residue that the information contained in the subbands is suitable for information processing [14].

B. Support vector machine (SVM)

Support vector machine (SVM) is a set of supervised learning methods used for classification and regression. Vepnik and Chervonenkis introduced SVM based on statistical learning theory. The support vector machine is essentially a mathematical algorithm for maximizing the mathematical function according to the given dataset. It is known for its success in recognizing handwritten letters that can be compared to complex neural networks [15]. Indeed, the SVM classification model employs the use of the optimal separation rule, where the input prototypes are transferred into a high-dimensional features space by employing a nonlinear mapping method. This rule selects an optimal separating hyperplane with the maximum distance between linearly separable classes. The concept of the optimal separating hyperplane is derived from the structural risk minimization theory, based on minimizing the empirical risk of the learning machine or on the defined upper bound on the error [21].

1) Linear classifier

A classifier is often defined as $f(x): R^d \to R$. In a 2 classes problem, the data belongs to the positive class if $f(x) \geq 0$; else it belongs to the negative class.

A linear function is a function that is defined from the linear combination of input x as (1):

$$f(x) = \sum w_{ij} x_j + b = w^T x + b \tag{1}$$

A set of points (x,y) where $y_i \in \{-1,1\}$ are linearly separable if a linear classifier exists which satisfies (2):

$$y_i f(x) \geq 0 \tag{2}$$

2) Hyperplane

A hyperplane is a concept in geometry, a generalization of a plane in a number of different dimensions. A hyperplane defines a k-dimensional sub-dimension of n-dimensional space that n>k. For example, a line is a one-dimensional hyper-plane in a space with any number of dimensions.

In three dimensions, the plane is a hyperplane in a two-dimensional space and is thus defined for spaces with higher dimensions. In SVM, any data is seen as a P-vector (a list of P-numbers) and we want to know if such points can be separated by a P-1-hyper plane. This is called linear classification. There are many hyperplanes that can separate the data. What distinguishes SVM from other classifiers is how to select the hyperplane. In SVM, maximizing the margin between the two sets is considered; If such a hyperplane exists, it is known as a maximum-margin hyperplane. The maximum margin and this concept visually are shown in Figure 2 [16]. In fact, the hyperplane was applied to separate the points having the label of +1 from those with the label of -1.

The decision-making function for data separation is determined by the support vectors (the closest training data to the classifier hyperplane). In fact, the optimal hyperplane in SVM is the separators between the supporting vectors. This

algorithm will have good generalization power if used properly by SVM. Also, it can work well in high dimensions and avoid overfitting. Additionally, due to the use of support vectors instead of whole data, this algorithm also performs data processing. As seen in the simplest case, the data can be separated by a straight line; But many real datasets are not so easily separable as when the database contains noise and errors. In this case, too, SVM provides nonlinear separator expansion and kernel function definition (for higher dimensions) and other solutions capable of accessing this type of data. The support vector machine analysis method is a new, fast, cheap, and non-destructive tool [17].

C. Multi-layer perceptron neural network (MLP NN)

In many complex mathematical problems that lead to the solution of nonlinear complex equations, a multilayer perceptron neural network can be used simply by defining appropriate weights and functions. Different activation functions are used according to the style of the problem in neurons. In this type of network, an input layer is used to apply the problem inputs of a hidden layer and an output layer, which ultimately provide the problem results. The nodes that are present at the input are sensory neurons and the nodes at the output are posterior neurons. There are also hidden neurons in the hidden layers. Training of such networks is usually done by the error backpropagation method [18]. The general topology of a multilayer perceptron network is shown in Figure 3.

Multi-layer perceptron networks can be constructed and used with any number of layers, but the theorem we accept here without proof states that a three-layer perceptron network can separate any type of space. This theorem, called the Kolmogorov theorem, represents a very important concept that can be used to construct neural networks. A special type of multi-layer network was called single-layer perceptron (SLP). This network as shown in figure 4, consists of an input layer and an output layer [19].

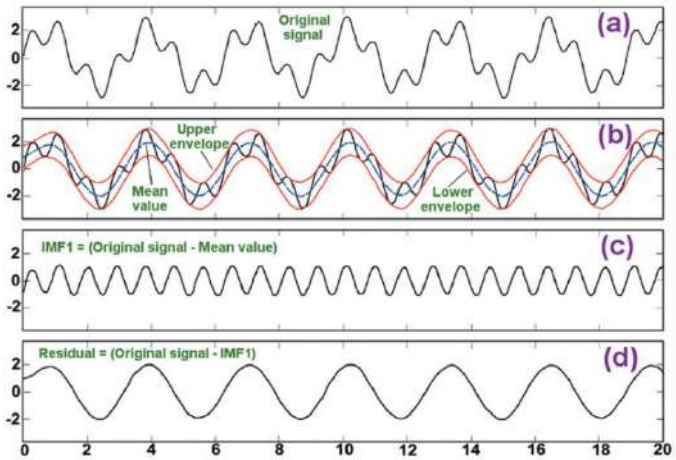


Figure 1. Procedure for applying EMD to a signal (a: The sum of the two signals. b: The top and bottom signal envelopes c: The first IMF d: The first residue of the original signal) [22].

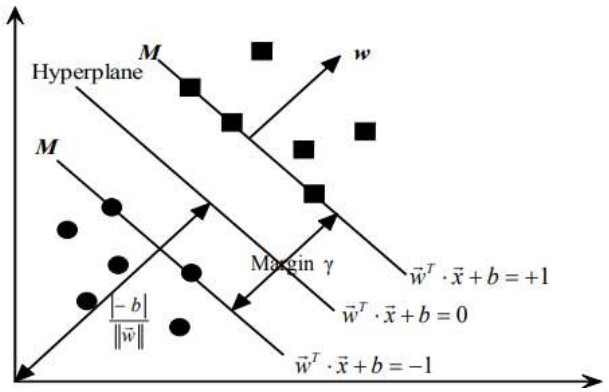


Figure 2. Hyper-plane and margin, circular dots, and square dots represent samples of class -1 and class +1 [16].

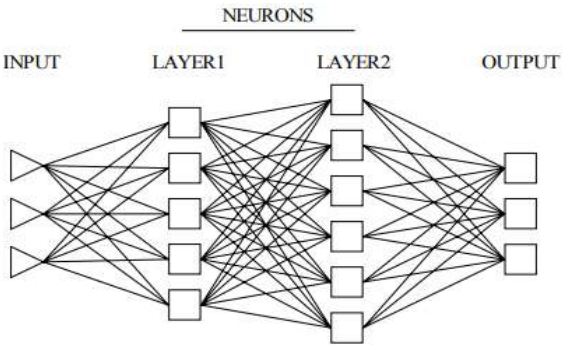


Figure 3. The general structure of Multi-layer perceptron [16].

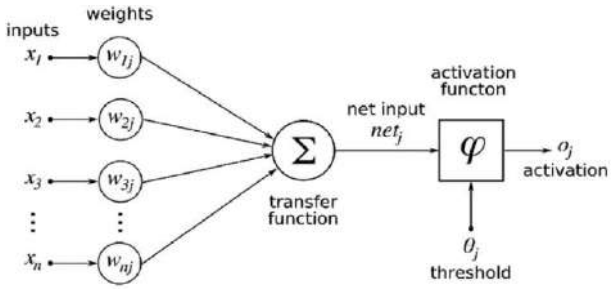


Figure 4. The single-layer perceptron topology [21].

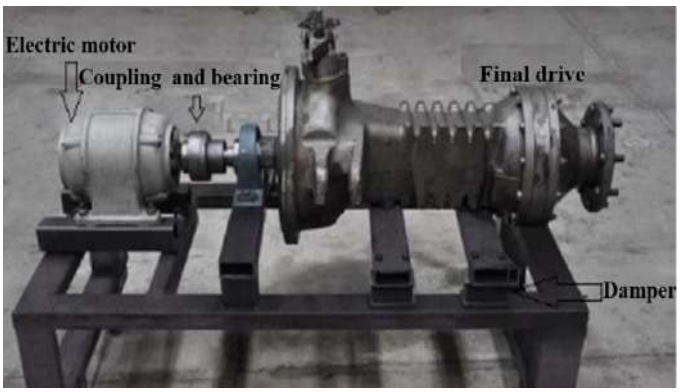


Figure 5. Experiment platform [23].

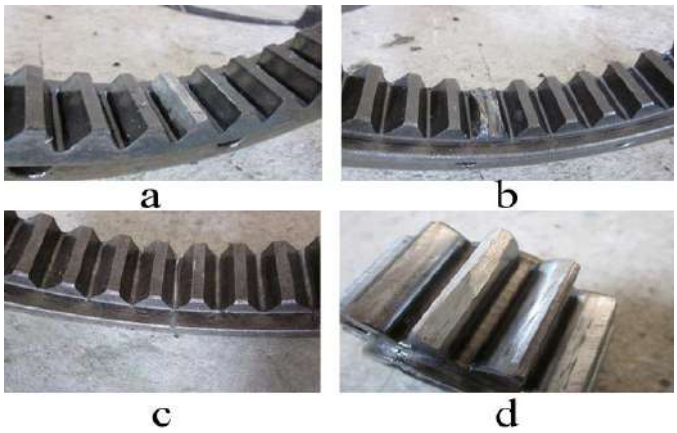


Figure 6. Different fault conditions in the planetary gearbox (a: Abrasion inside the gear. b: Fracture of the tooth inside the gear. c: Cracks in the teeth inside the gear. d: Planetary tooth wear) [23].

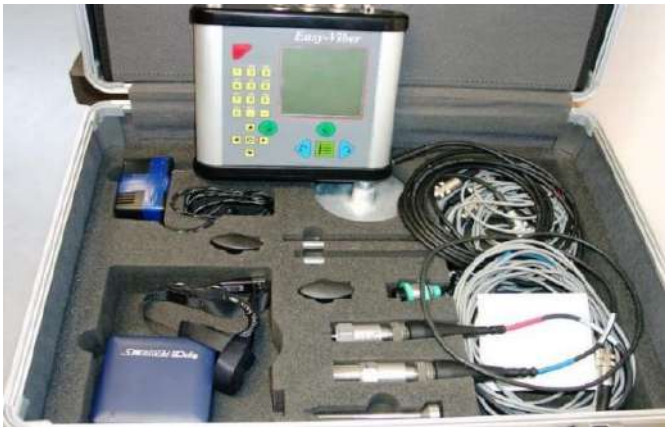


Figure 7. Data collection device and sensor.

III. EXPERIMENT PLATFORM

In order to collect vibration data from the final reducing function in healthy condition and gear failure, the seedling bed given in Figure 5 was prepared with the main components of the bed including the electric motor, pulley, coupling, chassis, and speed changer. Figure 6 shows the different cases of final reduction faults investigated in this study. The faults investigated in this research are evident in the inverter gear and other devices. The results obtained from the enthusiasm of the final reduction repairers showed that fractures, cracks, and abrasions in the gears and planetary gear tooth wear of the common planetary faults occurring in the final gear. 0.5 mm abrasion was created on both sides of the in-gear and planetary teeth, but to simulate fractures and cracks in the gears, these two types of gears were obtained from repair shops.

A. Tools for recording data

In this study, vibration data from different modes of working planetary gearboxes were measured and recorded. For this purpose, Easy-Viber has been used. The Easy-Viber device is made by VMI, Sweden. This device has a piezoelectric accelerometer of type VMI-102 and a tachometer to measure and record it. In general, vibrations were measured and recorded on the body of the drive gear and near the bearing at the rotational speed of the input to the final drive at 300 rpm. The sensor was attached to the body by a magnetic probe. For each case, 30 vibration signals were recorded. The sampling frequency of the device was 8192 Hz and the length of each signal was two seconds. Figure 7 shows the data acquisition set.

B. Signal processing

Although vibration data carries very important and useful information about the condition of the machine, they not only include the fault signal but also have a lot of noise. The presence of these noises in the signals of the time domain directly causes serious problems. Solving this problem requires that the signals be transmitted from the time domain to the time-frequency domain so that in addition to eliminating noise, more useful information can be sent to them. In the present study, the empirical mode decomposition method is used to process vibration signals.

C. Feature extraction

It is not possible to use the output of the signal processing method directly in existing systems. For this reason, features of decomposed subbands must be extracted using functions. Statistical features and some of the functions used in monitoring and positioning that were used to extract the features in this study include mean, harmonic mean, square root mean geometric mean, standard deviation, signal peak amplitude value, signal sharpness factor, Variance, skewness, and kurtosis. more information on features can be found on [20].

D. Integrated model of fault diagnosis

In this research, an integrated model is used to diagnose and classify the faults of the planetary gearbox. The proposed model is a combination of a support vector machine and an artificial neural network, the flowchart of which is shown in Figure 8. 70% of the data will be used for training and the remaining 30% for testing. First, an artificial neural network model was created using the training data. After training the neural network model, the output layer of the neural network was removed and the shortened neural network output was directly considered as the input of the support vector machine, and the support vector machine was trained. Finally, the integrated neural network-support vector machine model was used to diagnose the type of fault.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

Figure 9 shows the time signal for different fault conditions in the gearbox. A general examination of the signals showed that their vibration behavior would change when a fault occurred in the gearbox. When faults occur, not only the amplitude of the vibration changes but also their oscillating frequencies are changed, but the appearance in the time domain cannot be a powerful tool in troubleshooting. For this reason, all vibration signals were processed by the method of EMD, and each vibration signal was decomposed into six subbands or IMF, and finally, 10 statistical features were extracted from each subband. In total, the results showed that 60 statistical features were extracted from each vibration signal, which was used to create an integrated model of artificial intelligence.

In this study, a total of 30 vibration signals were studied in each case. This number of samples was divided into two parts: 21 samples were used for training and classification structure

and 9 samples were used in testing the classification. A total of 60 features were used as inputs to the neural network classifier. The results obtained in the classification of faults by ANN and SVM are represented as confusion matrices in Tables 1 and 2, respectively. The confusion matrix obtained from the training and testing of the integrated model is given in Tables 3 and 4. Moreover, the parameters for evaluating the efficiency of the artificial neural network in diagnosing the faults of the planetary system are presented in Tables 1, 2, 3, and 4. The results showed that as the MLP NN model gets into local minima, is not a suitable solution for this fault classification task. Furthermore, not only the SVM model has higher accuracy but also it has solved the mentioned problem. Additionally, the proposed integrated model can be trained with 97.14% accuracy. The sensitivity of the integrated model trained in the diagnosis of gearbox health status data was 100%. The results of testing the model showed that the proposed method has an overall accuracy of 91.11% in fault diagnosis. This model was able to classify five different gearbox states with high accuracy. It can also be inferred that the distribution of the collected data is such that we will see a low performance. As a result, integrated algorithms should be used to have more accuracy as suggested.

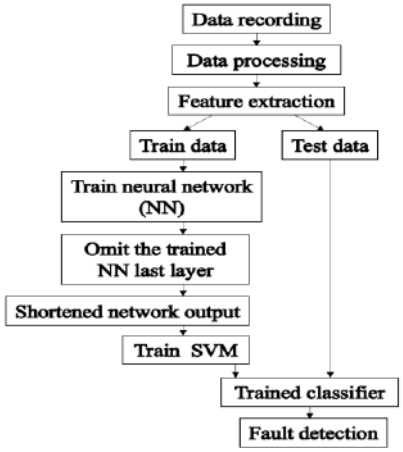


Figure 8. Integrated model for diagnosing and classifying the type of planetary gearbox fault.

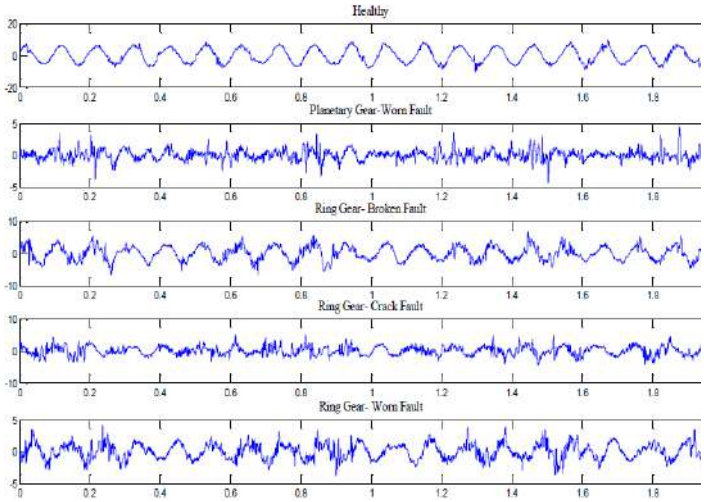


Figure 9. Vibration signals of the planetary system in order from top to bottom: healthy state, abrasive planets, peripheral fracture, peripheral cracking, and peripheral abrasion.

V. CONCLUSION

Today, the ever-increasing need for greater accuracy, reliability and security in modern industries has led to the widespread use of vibration analysis methods in troubleshooting industrial equipment. In this research, an effective and powerful method for accurate fault diagnostics of planetary gearboxes based on vibration data is presented. For this purpose, first vibration signals were studied in different states of the gears and then processed by empirical mode decomposition method and each signal was decomposed into six subbands. 10 features were extracted from each subband. Finally, 60 features were extracted from each signal. Using these features, the integrated model of an artificial neural network-support vector machine was taught. The results showed that the integrated model was trained with an accuracy of 97.14%, which finally this model was able to detect gearbox faults for the test data with an accuracy of 91.11%. Moreover, the integrated model has higher accuracy than MLP and SVM methods and can be presented as an efficient method without getting into local minima.

TABLE I. MLP MODEL CONFUSION MATRIX IN PLANETARY GEARBOX FAULT DETECTION FOR TEST DATA. (HEALTH (H), RING BROKEN (RB), RING WORN (RW), RING CRACK (RC), SUN WORN (SW))

Real / Predicted	H	RB	RW	RC	SW	Sensitivity(%)
H	0	10	0	0	0	39
RB	0	7	0	0	0	35.66
RW	0	9	0	0	0	52
RC	0	6	0	7	0	50.95
SW	0	0	0	0	6	93.82
Total Accuracy						49.13%

TABLE II. SVM MODEL CONFUSION MATRIX IN PLANETARY GEARBOX FAULT DETECTION FOR TEST DATA. (HEALTH (H), RING BROKEN (RB), RING WORN (RW), RING CRACK (RC), SUN WORN (SW))

Real / Predicted	H	RB	RW	RC	SW	Sensitivity(%)
H	5	4	0	0	0	47.82
RB	1	5	0	0	0	53.06
RW	0	0	12	0	0	100
RC	0	0	0	9	0	100
SW	0	0	0	0	9	100
Total Accuracy						79.33%

TABLE III. INTEGRATED MODEL CONFUSION MATRIX IN PLANETARY GEARBOX FAULT DETECTION FOR TRAINING DATA. (HEALTH (H), RING BROKEN (RB), RING WORN (RW), RING CRACK (RC), SUN WORN (SW))

Real / Predicted	H	RB	RW	RC	SW	Sensitivity(%)
H	21	0	0	0	0	100
RB	0	21	0	0	0	100
RW	0	0	19	1	1	90.48
RC	0	0	0	21	0	100
SW	0	1	1	0	20	95.24
Total Accuracy						97.14 %

TABLE IV. INTEGRATED MODEL CONFUSION MATRIX IN PLANETARY GEARBOX FAULT DETECTION FOR TEST DATA. (HEALTH (H), RING BROKEN (RB), RING WORN (RW), RING CRACK (RC), SUN WORN (SW))

Real / Predicted	H	RB	RW	RC	SW	Sensitivity(%)
H	9	0	0	0	0	100
RB	0	8	0	1	0	88.89
RW	0	1	8	0	0	88.89
RC	0	0	0	9	0	100
SW	0	0	0	2	7	77.78
Total Accuracy						91.11 %

REFERENCES

[1] Salameh, J.P., et al., Gearbox condition monitoring in wind turbines: A review. Mechanical Systems and Signal Processing, 2018. 111: p. 251-264.

[2] Jing, L., et al., A convolutional neural network based feature learning and fault diagnosis method for the condition monitoring of gearbox. Measurement, 2017. 111: p. 1-10.

[3] Zhou, P., et al., Sinusoidal FM patterns of fault-related vibration signals for planetary gearbox fault detection under non-stationary conditions. Mechanical Systems and Signal Processing, 2021. 155: p. 107623.

[4] Yao, J., et al., Fault detection of complex planetary gearbox using acoustic signals. Measurement, 2021. 178: p. 109428.

[5] Rgeai, M.N., Helical gearbox fault detection using motor current signature analysis. 2007: The University of Manchester (United Kingdom).

[6] Vamsi, I., G. Sabareesh, and P. Penumakala, Comparison of condition monitoring techniques in assessing fault severity for a wind turbine gearbox under non-stationary loading. Mechanical Systems and Signal Processing, 2019. 124: p. 1-20.

[7] Roshanmanesh, S., F. Hayati, and M. Papaelias, Utilisation of ensemble empirical mode decomposition in conjunction with cyclostationary technique for wind turbine gearbox fault detection. Applied Sciences, 2020. 10(9): p. 3334.

[8] Reza Asadi Asad Abad, M., et al., Discrete wavelet transform and artificial neural network for gearbox fault detection based on acoustic signals. Journal of Vibroengineering, 2013. 15(1): p. 459-463.

[9] Yang, S., W. Li, and C. Wang. The intelligent fault diagnosis of wind turbine gearbox based on artificial neural network. in 2008 international conference on condition monitoring and diagnosis. 2008. IEEE.

[10] Chen, P., et al., An automatic speed adaption neural network model for planetary gearbox fault diagnosis. Measurement, 2021. 171: p. 108784.

[11] Ebrahimi, E., A method based on support vector machine and vibration analysis for fault detection in bevel gears (Case study: Differential). Insight-Non-Destructive Testing and Condition Monitoring, 2019. 61)5(: p. 279-286.

[12] Liu, Y., W. Ni, and Z. Ge, Fuzzy decision fusion system for fault classification with analytic hierarchy process approach. Chemometrics and Intelligent Laboratory Systems, 2017. 166: p. 61-68.

[13] Omrani, T., et al. Fusion of ANN and SVM classifiers for network attack detection. in 2017 18th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA). 2017. pppp.

[14] Rostami, J., J. Chen, and P.W. Tse, A signal processing approach with a smooth empirical mode decomposition to reveal hidden trace of corrosion in highly contaminated guided wave signals for concretecovered pipes. Sensors, 2017. 17(2): p. 302.

[15] Ma, Y. and G. Guo, Support vector machines applications. Vol. 649. 2014: Springer.

[16] Kianmehr, K., Zhang, H., Nikolov, K., & Alhajj, R. (2011). Combining Neural Network and Support Vector Machine Into Integrated Approach for Biodata Mining. 182–187.

[17] Jakkula, V., Tutorial on support vector machine (svm). School of EECS, Washington State University, 2006. 37.

[18] Mahmud, M.N., et al., A robust transmission line fault classification scheme using classdependent feature and 2-Tier multilayer perceptron network. Electrical Engineering, 2018. 100(2): p. 607-623.

[19] Mishra, K.M. and K.J. Huhtala. Fault detection of elevator systems using multilayer perceptron neural network. in 2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). 2019. IEEE.

[20] Farokhzad, S., et al., 897. Artificial neural network based classification of faults in centrifugal water pump. Journal of Vibroengineering, 2012. 14(4).

[21] Abbas, A. K., Al-haideri, N. A., & Bashikh, A. A. (2019). Implementing artificial neural networks and support vector machines to predict lost circulation. Egyptian Journal of Petroleum, 28(4), 339–347.

[22] Hassan, H. H. (2005). Empirical Mode Decomposition (EMD) of potential field data: Airborne gravity data as an example. Society of Exploration Geophysicists - 75th SEG International Exposition and Annual Meeting, SEG 2005, Figure 1, 704–706.

[23] Khazaee, M., Ahmadi, H., Omid, M., Moosavian, A., & Khazaee, M. (2012). Vibration condition monitoring of planetary gears based on decision level data fusion using Dempster- Shafer theory of evidence. Journal of Vibroengineering, 14(2), 838–851.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Intelligent Vibration-based Anomaly Detection for Electric Motor Condition Monitoring

Tuan A. Z. Rahman, Leong Wen Chek and Nordin Ramli

IoT Systems Laboratory, MIMOS Berhad 57000 Kuala Lumpur, Malaysia
zahidi.rahman@mimos.my

***Abstract*— The health state of rotating machinery can be represented by its vibration signal. With the help of machine learning approaches, the condition of rotating machinery such as electric motors can be diagnosed accurately. Thus, equipment down-time can be minimized and catastrophic accidents can be avoided. This paper presents an intelligent anomaly detection method for electric motors based on vibration signals. Due to lack of damage conditions information, only the normal condition data were employed to generate an unsupervised learning model for two different types of motor within the same class, which are the new laboratory motor and old industrial ones. The performance of the model to detect anomalies for both motors was studied, extensively. The results show that the model generated possesses the highest capability to detect anomalies for both motors as the normalized and mapped features were used.**

***Index Terms*—Anomaly detection, Benchmarking, Condition monitoring, Machine learning, Population-based**

I. Introduction

Nowadays, condition-based monitoring (CBM) and predictive maintenance have become very popular areas of research due to high demand in development of IoT-based smart manufacturing for Industrial Revolution 4.0 (IR4.0) purpose. It is observed that CBM and predictive maintenance approaches have massive market opportunities. And with the help of machine learning, the innovative solution to CBM and predictive maintenance can be implemented [1]. Previously, most of the machine learning models for CBM were trained based on single equipments vibration signal dataset. However, when it comes to implement the developed model onto other set of equipment, the prediction accuracy of the model dropped as different dataset is used. Thus, the objective of this study is to develop a machine learning model for monitoring two different electric motors within similar class based on their vibration signal. Features mapping technique is proposed so that the model's input is within a similar range.

The rest of this paper is organized as follows: Section 2 discusses regarding previous related works on population-based monitoring using data mining methods. The experimental data collection and development of proposed anomaly detection algorithm are explained in Section 3. Section 4 presents the main results of developed algorithm and comparative study on its performance. Finally, Section 5 concludes the paper and gives possible future directions for research.

II. Related Work

Recently literature surveys in [2-4] highlighted the challenges and opportunities in order to develop robust predictive maintenance techniques based on machine learning especially for rotational equipment such as bearing, motor, gear, pump and etc.. There are still many rooms that can be explored in this area in order to improve the machine learning model's prediction accuracy and increase the versatility of the proposed predictive maintenance approaches. Several previous works have been studied on the population-based condition monitoring of rotational equipments and structures.

Hendrickx *et al.* [5] proposed an unsupervised, generic, anomaly detection framework for fleet-based condition monitoring by using generic building blocks. They highlighted three advantages of the proposed approach such as no need for historical dataset, combining domain expertise by user-defined comparison measures and easy interpretability. The proposed approach was implemented to detect a voltage unbalance by means of electrical and vibration signatures of electric motors.

Bull *et al.* [6] proposed a framework of population Structural Health Monitoring (SHM) to model a population of nominally-identical systems by observing to both model and transfer the missing information using data collected from groups of similar structures. The population-based SHM strategy defines a general model is used to monitor a homogeneous group of systems. The framework demonstrated through applications to a simulated population and signals recorded from an operational wind farm.

Based on these literature studies, it can be said that population-based CBM for electric motors will be improvement from the traditional monitoring techniques as a newly system was upgraded to monitor a homogeneous group of motors with similar characteristics.

III. Methodology

In this section, the proposed development of intelligent anomaly detection is elaborated in details. Flowchart of the proposed development of anomaly detection algorithm was illustrated as in Fig. 1. There are several stages to be complete in order to develop a robust anomaly detection for the electric motors including experimental data collection, data pre-processing/cleansing, data mining techniques selection, and model's training and validation.

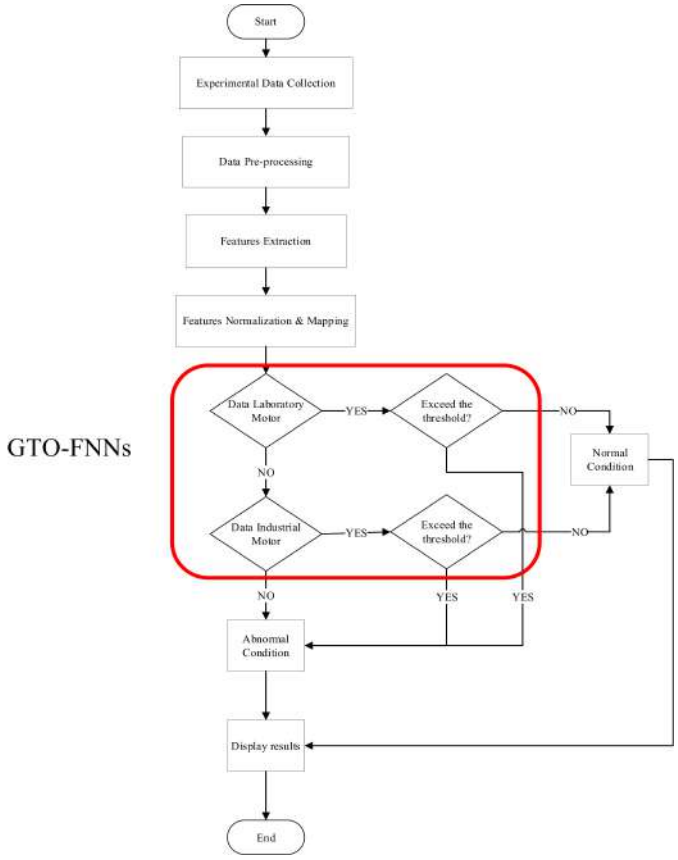


Fig. 1. Flowchart of the developed algorithm

A. Data Collection

The experimental data collection was done using Wiserot system developed by Fuji Electric. Wireless accelerometers were attached to the motors that working in normal operational conditions in laboratory and industry. The new electric motor in laboratory was linked to a dynamotor system that drive a constant workloads through a pulley. While, the old motor in the industry running a mincer to produce chilli sauce. Both motors were identified within similar class which is Class 2. Fig. 2 shown the experimental set-up to collect the vibration data for both motors. The motors were running with low speed of 1500 rpm. Default sampling frequency of 2048 Hz was employed to measure the vibration in velocity unit (mm/s). The exemplar of collected vibration data from both motors was plotted in Fig.3.

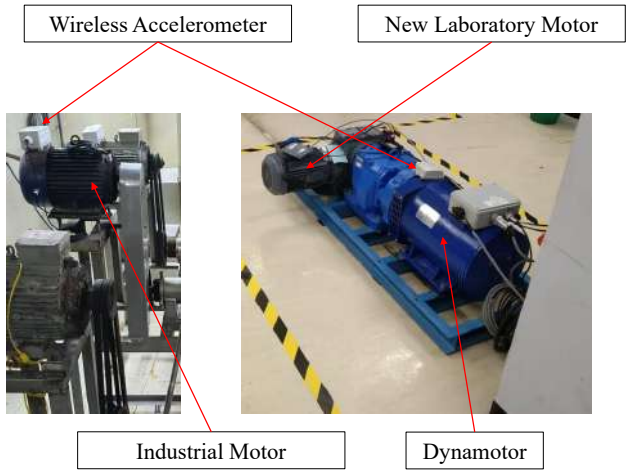


Fig. 2. Experimental set up for data collection

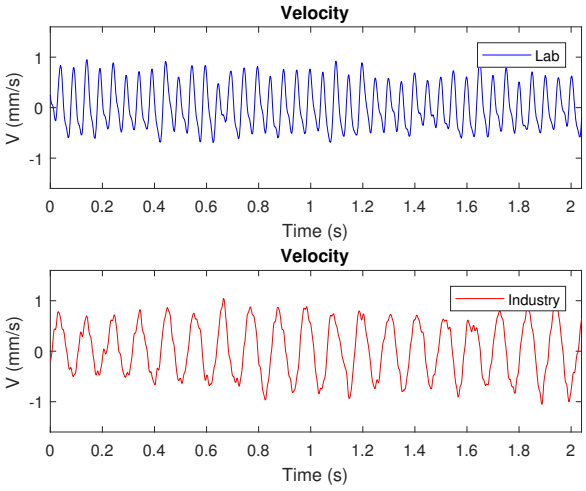


Fig. 3. The exemplar of operation vibration signals for motors

B. Model Development

The development of the anomaly detection model using machine learning algorithms and their performance criteria is presents in this subsection. The collected vibration data were pre-processed which including data range selection and normalization before performing the features extraction phase. Then, the dataset were exported as images into an open source Python-based software named Orange data mining. The features were extracted from each images using image embedded technique known as VGG-19, an embedder proposed by Visual Geometry Group from the University of Oxford [7].

A total number of 330 observations from vibration signals of the two motors were computed into high dimensional features of 4096 instances. Then, six types of machine learning (ML) algorithms were employed to classify the source of vibration signal either from laboratory or industry motor by using their default settings within the Orange software. The performance of these ML algorithms were calculated from the results of confusion matrix table based on Eq.s 1-4.

$$CA = (TP + TN)/(TP + TN + FP + FN) \quad (1)$$

$$Precision = TP/(TP + FP) \quad (2)$$

$$Recall = TP/(TP + FN) \quad (3)$$

$$F1 = 2TP/(2TP + FP + FN) \quad (4)$$

where, *CA* is classification accuracy, *TP* is true positive, *TN* is true negative, *FP* is false positive and *FN* is false negative. Other performance criteria also computed are *Precision*, *Recall* (sensitivity) and *F1* score which is combination of *Precision* and *Recall* in a single metric.

Later, Principle Component Analysis (PCA) technique was used to visualise the dataset and to reduce the features dimensionality. The first two principle components (PCs) were selected and then normalised between -1 and 1 for both motor's dataset. All the calculation process done within the MATLAB environment.

Fig. 4 shows the two layers of feedforward neural networks (FNNs) model employed to develop anomaly detection for electric motor in the study. The input to the model are results from features mapping of two first principle components which are PC1 and PC2 into [-1, 1] dimension space. The architecture of networks 4-10-1 was used in the study as suggested in [8] with 10 neurons in the hidden layer. Output layer will generate result of 0 and 1 to indicate normal and faulty conditions, respectively.

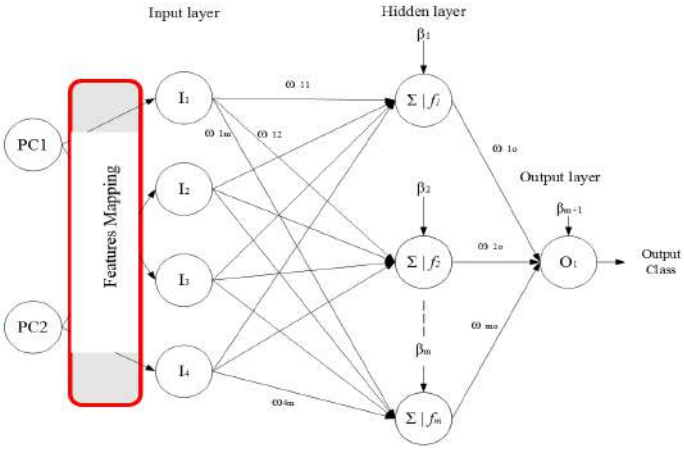


Fig. 4. Two layers of neural networks model for anomaly detection

The model was trained using a new metaheuristic algorithm named as artificial gorilla troops optimizer (GTO) proposed by Abdollahzadeh *et al.*. The details regarding the GTO algorithm can be find in [9]. Mean squared of error (MSE) between actual and predicted outputs is selected as objective function. A total numbers of 61 weight and bias coefficients for neurons were computed within range of [-10, 10] using population size of 30 searching agents. The default control parameter settings for GTO algorithm was applied to optimize the FNNs model as suggested in [9] with *p*=0.03, *Beta*=3 and *w*=0.8.

In validation phase, faulty condition dataset was created artificially to test the accuracy of developed FNNs model. Four conditions were considered with different locations within the PC features space. The validation process was repeat for 1000 times and the average accuracy was calculated for each condition.

IV. Results and Discussion

An AMD desktop (64-bit Ryzen-5) with 16 GB RAM was used for computational power in Python 3.8 and then later in MATLAB environments. Fig. 5 visualises the condition of dataset before and after feature mapping using PCA method. Based on the figure, it is clearly can be see that the readings from both motors are separable as image embedded technique via deep learning was used to extracted the features. After normalisation process, the dataset were mapped into same range of [-1, 1] for both motor dataset. Again, the classification results of six ML algorithms in Table 1 shows high accuracy level, also indicated that the dataset from new laboratory and old industry motors can clearly be differentiated especially when using *k*-Nearest Neighbour (*k*-NN) and Logistic Regression algorithms.

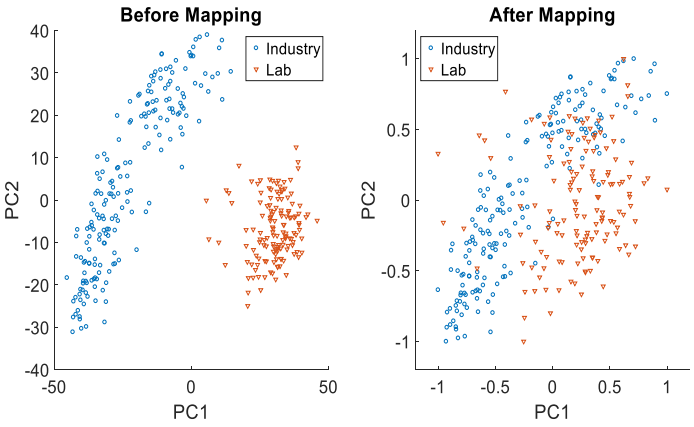


Fig. 5. The visualization of dataset using PCA technique before and after features mapping

TABLE I
THE PERFORMANCE OF SIX MACHINE LEARNING ALGORITHMS

Algorithm	Performance Evaluation			
	CA	***F1***	***Precision***	***Recall***
kNN	1.000	1.000	1.000	1.000
Logistic Regression	1.000	1.000	1.000	1.000
Random Forest	0.997	0.997	0.997	0.997
Neural Network	0.997	0.997	0.997	0.997
AdaBoost	0.991	0.991	0.991	0.991
SVM	0.976	0.976	0.977	0.976

The convergence plot of GTO algorithm in training the FNNs model over the course of 5000 number of function evaluations (NFEs) is shown in Fig. 6. The best optimized-MSE achieved was 1.52×10^{-4}. The GTO algorithm possess a great capability to find an appropriate optimum MSE value within a low computational time.

Fig.s 7 and 8 show the performance of the developed FNNs model to predict the condition of electric motor based on

different locations of artificial faulty dataset in the first two PCs coordinate grid. A total number of 600 faulty dataset were generated in each scenarios using Gaussian random number generator. The first scenario was located outside of normal condition regions for both motors at location [-80, 20] in the PC1 vs PC2 grid while in the second scenario was located between normal condition regions at location [-10, -60].

In both scenarios, the developed model yielded highly prediction accuracies with average of 99.9% to indicate that the tested dataset were in faulty conditions. Subsequently, the distance between the normal and faulty regions were reduced that lead to decrease in average accuracy of the model as the model mistaken the faulty conditions with the normal ones (as in Fig.s 7(c) and 8(c)). The centroid of artificial faulty dataset were transposed to [-60, 20] and [0, -40] for the first and second scenario, respectively. In Fig.s 7(b), 7(d), 8(b) and 8(d), the predicted and rounded output values from the model were depicted.

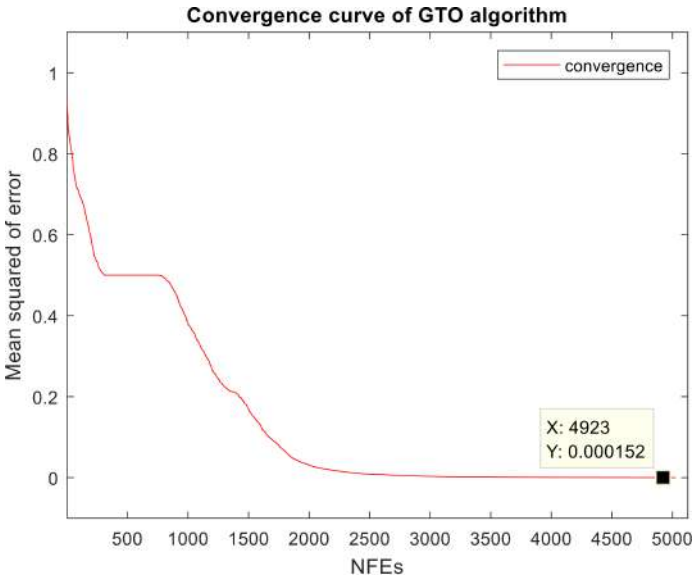


Fig. 6. The convergence curve of GTO algorithm in training FNNs

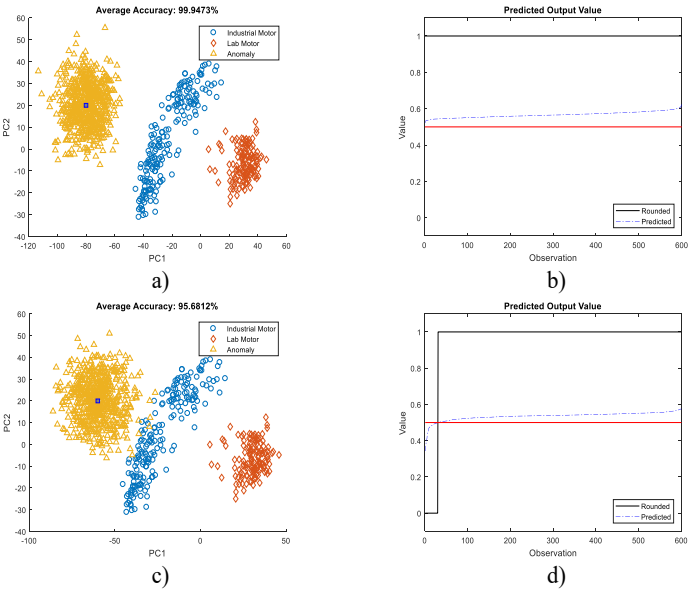


Fig. 7. Prediction of the developed FNNs model in the first scenario

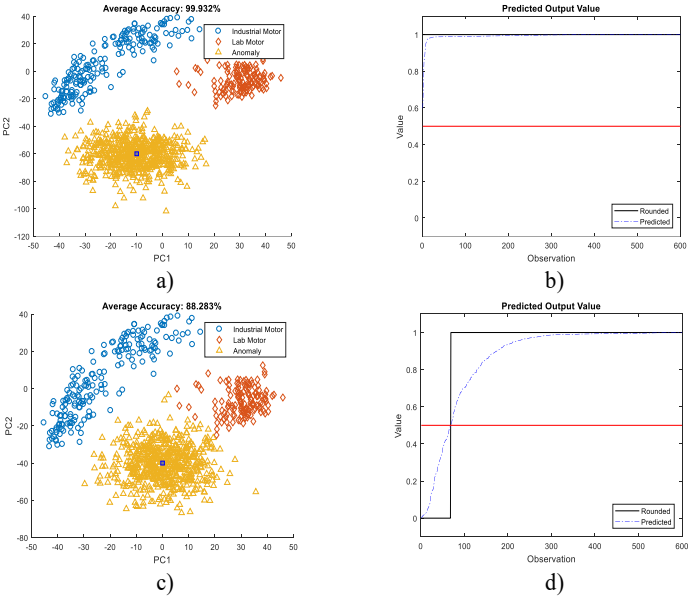


Fig. 8. Prediction of the developed FNNs model in the second scenario

V. Conclusion

In this work, a new intelligent vibration-based anomaly detection algorithm for condition monitoring of two electric motors has been proposed using metaheuristic-optimized Feed-forward Neural Networks (FNNs) model. In the first stage, the experimental vibration motor dataset was pre-processed in Orange data mining software in order to understand their characteristics and to distinguish between their sources. After that, the two layers NNs was trained using the GTO metaheuristic algorithm within the MATLAB environment based on normalised features dataset of both electric motors. As a results, the developed model displayed high prediction accuracy when tested with artificial faulty conditions. The selection of features extraction method plays important roles to improve the prediction accuracy. The future works include the validation process of developed model using actual faulty conditions and online IoT-based anomaly detection for monitoring a fleet of electric motors.

Acknowledgment

The authors would like to show gratitude to MIMOS Berhad for the funding to perform this research. The authors would like to thank the reviewers for their valuable feedback and comments.

References

[1] Z. M. Çınar, A. Abdussalam Nuhu, Q. Zeeshan, O. Korhan, M. Asmael and B. Safaei, "Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0," Sustainability, vol. 12(19):8211, pp. 1-42, October 2020.

[2] A. Rai and S. H. Upadhyay, "A review on signal processing techniques utilized in the fault diagnosis of rolling

element bearings," Tribol. Int., vol. 96, pp. 289-306, April 2016.

[3] S. R. Saufi, Z. A. B. Ahmad, M. S. Leong and M. H. Lim, "Challenges and opportunities of deep learning models for machinery fault detection and diagnosis: A review," IEEE Access, vol. 7, pp. 122644-122662, August 2019.

[4] Z. Zhao, J. Wu, T. Li, C. Sun, R. Yan and X. Chen, "Challenges and opportunities of AI-enabled monitoring, diagnosis and prognosis: A review," Chin. J. Mech. Eng., vol. 34(56), pp. 1-29, June 2021.

[5] K. Hendrickx, W. Meert, Y. Mollet, J. Gyselinck, B. Cornelis, K. Gryllias and J. Davis, "A general anomaly detection framework for fleet-based condition monitoring of machines," Mech. Syst. Signal. Process., vol. 139(106585), pp. 1-21, May 2020.

[6] L. A. Bull et al., "Foundations of population-based SHM, Part I: Homogeneous populations and forms," Mech. Syst. Signal. Process., vol. 148(107141), pp. 1-19, February 2021.

[7] K. Simonyan and A. Zisserman,"Very deep convolutional networks for large-scale image recognition," in The 3rd International Conference on Learning Representations (ICLR2015) San Diego, CA, USA, May 7-9, 2015.

[8] S. Mirjalili, S. Z. M. Hashim and H. M. Sardroudi, "Training feedforward neural networks using hybrid particle swarm optimization and gravitational search algorithm," Appl. Math. Comput., vol. 218(22), pp. 11125-11137, July 2012.

[9] B. Abdollahzadeh, F. Soleimanian-Gharehchopogh and S. Mirjalili, "Artificial gorilla troops optimizer: A new nature-inspired metaheuristic algorithm for global optimization problems," Int. J. Intell. Syst., vol. 36, pp. 5887-5958, July 2021.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Intuitionistic Fuzzy Multiset finite Subautomata

1st Marzieh Shamsizadeh, 2nd Mohammad Mehdi Zahedi and 3rd Khadijeh Abolpour

1st Department of Mathematics, Behbahan Khatam Alanbia University of Technology, Khouzestan, Iran, shamsizadeh.m@gmail.com
2nd Department of Mathematics, Graduate University of Advanced Technology, Kerman, Iran, zahedi_mm@kgut.ac.ir
3rd Department of Mathematics, Shiraz Branch, Islamic Azad University, Shiraz, Iran, abolpor_kh@yahoo.com

***Abstract*— In this note, we first present the notion of intuitionistic fuzzy multiset finite automata. After that, we present the notion of successor with threshold (a, b) for every state q, where $a, b \in [0, 1]$ and $0 \leq a + b \leq 1$. Also, by considering the concept of successor, we introduce the notion of intuitionistic fuzzy multiset finite subautomata (IFMFSA). Furthermore, we show that the union and intersection intuitionistic fuzzy multiset finite subautomata is an IFMFSA, too.**

***Index Terms*—Automata; Intuitionistic; Multiset Automata; Fuzzy Automata**

I. Introduction

The notion of fuzzy set as a method for representing uncertainty has been introduced by Zadeh in 1965 [15]. Wee [14] and Santos [8] have introduced the idea of fuzzy automata.

The intuitionistic fuzzy sets introduced by Atanassov [1] have been found to be highly useful to deal with vagueness. Using the notion of intuitionistic fuzzy sets, Jun [6] introduced the notion of intuitionistic fuzzy finite state machines as a generalization of fuzzy finite state machines. Based on the papers [6, 7], Zhang and Li [16] discussed intuitionistic fuzzy recognizers and intuitionistic fuzzy finite automata. In 2015, Shamsizadeh and Zahedi gave the notion of max-min intuitionistic general fuzzy automaton [9–11].

A multiset, which is a collection of elements where elements can occur several times, is a generalization of a set [3]. In recent years, multiset processing has emerged substantially in different fields of mathematics, computer science, biology and biochemistry. Csuhajvarjú Martin-Vide and Mitrana [5] have proposed multiset automata and multiset grammars have been suggested in[4], as they investigated some of the properties of Mealy multiset automata.

Now, in this paper we present the notion of intuitionistic fuzzy multiset finite automata. Also, we present the notion of successor with threshold (a, b) for every state q, where $a, b \in [0, 1]$ and $0 \leq a + b \leq 1$. After that, by considering the concept successor we present the notion of intuitionistic fuzzy multiset finite subautomata.

II. Preliminaries

In this section, some concepts and definitions related to multisets and automata are introduced.

Definition 1. *[3, 5] If Σ is a finite alphabet, then $\alpha : \Sigma \to N$ is a multiset over Σ, where N denotes the set of natural numbers including 0. The α norm of Σ is defined by $|\alpha| = \sum_{a\in\Sigma} \alpha(a)$.*

We shall denote by $\Sigma^{\oplus}$, the set of all multiset over Σ. The multiset $0_\Sigma \in \Sigma^{\oplus}$ is defined by $0_\Sigma(a) = 0$, for every $a \in \Sigma$. For $b \in \Sigma$, we shall denote by $< b >$, a singleton multiset, and is defined by:

$$< b > (a) = \begin{cases} 1 & \text{if } b = a \\ 0 & \text{otherwise} \end{cases},$$

for every $a \in \Sigma$. For a given set A, let $\bar{A} = \{< a > \big| a \in A\}$.

For two multisets $\alpha, \beta \in \Sigma^{\oplus}$, the operations inclusion $\subseteq$, addition $\oplus$ and difference $\ominus$ are defined as follows:

1) $\alpha \subseteq \beta$ if $\alpha(a) \leq \beta(a)$,
2) $(\alpha \oplus \beta)(a) = \alpha(a) + \beta(a)$,
3) $(\alpha \ominus \beta)(a) = \max(0, \alpha(a) - \beta(a))$,

for every $\alpha \in \Sigma$. Furthermore, $\alpha \subset \beta$ if $\alpha \subseteq \beta$ and $\alpha \neq \beta$.

Clearly, $\Sigma^{\oplus}$ is a commutative monoid with identity element 0_Σ with respect to $\oplus$.

Definition 2. *[5] A multiset finite automata (MFA) is a 5-tuple $\mathcal{M} = (Q, \Sigma, \delta, q_0, F)$, where*

1. *Q and Σ are nonempty finite sets called the state-set and input-set,*
2. *$\delta : Q \times \Sigma^{\oplus} \to 2^Q$ is a map called transition map,*
3. *$q_0 \in Q$ is called the initial state,*
4. *$F \subseteq Q$ is called the set of final states.*

Now, we recall the following concepts of fuzzy multiset finite automaton (FMFA) from [13].

Definition 3. *A fuzzy multiset finite automaton (or FMFA, for short) is a 5-tuple $\mathcal{M} = (Q, \Sigma, \delta, \iota, \tau)$, where*

1. *Q and Σ are nonempty finite sets called the state-set and input-set, respectively,*

2. $\delta : Q \times \Sigma^{\oplus} \times Q \to [0,1]$ *is a map called fuzzy transition map,*
3. $\iota : Q \to [0,1]$ *is a map called the fuzzy set of initial states,*
4. $\tau : Q \to [0,1]$ *is a map called the fuzzy set of final states.*

A configuration of fuzzy multiset finite automaton $\mathcal{M}$ is a pair (q, β), where q and β denote current state and current multiset, respectively. The transition in a fuzzy multiset finite automaton is described with the help of configurations. The transition from configuration (q, β) leads to configuration (p, γ) with membership value $k \in [0,1]$ if there exists a multiset $\alpha \in \Sigma^{\oplus}$ with $\alpha \subseteq \beta, \delta(q, \alpha, p) = k$ and $\gamma = \beta \ominus \alpha$ and is denoted by $(q, \beta) \xrightarrow{k} (p, \gamma)$.

$\xrightarrow{k'}{}^{*}$ denote the reflexive and transitive closure of $\xrightarrow{k'}$, i. e., for $(q, \beta), (p, \gamma) \in Q \times \Sigma^{\oplus}$, $(q, \beta) \xrightarrow{k'}{}^{*} (p, \gamma)$ if for some $n \geq 0$, there exist $(n+1)$ states $q_0, ..., q_n$ and $(n+1)$ multisets $\beta_0, \beta_1, ..., \beta_n$ such that $p_0 = q, p_n = p, \beta_0 = \beta, \beta_n = \gamma$ and $(p_i, \beta_i) \xrightarrow{k_i} (p_{i+1}, \beta_{i+1})$, for every $i = 0, 1, ..., n-1$, where $k' = k_0 \wedge k_1 \wedge ... \wedge k_{n-1}$. Now, we define

$$\mu_M((q,\beta) \to^* (p,\gamma)) = \vee\{\mu_{\mathcal{M}}((q,\beta)) \to^* (r, \beta \ominus \alpha)) \wedge \mu_{\mathcal{M}}((r, \beta \ominus \alpha)) \to^* (p,\gamma)) \big| r \in Q, \alpha \in \Sigma^{\oplus}, \alpha \subseteq \beta\},$$

and

$$\mu_{\mathcal{M}}((q,\beta)) \to^* (p,\beta)) = \begin{cases} 1 & \text{if } q = p \\ 0 & \text{if } q \neq p \end{cases}. \quad (1)$$

In this note, consider $\Delta = \{\alpha \in \Sigma^{\oplus} \big| \delta(q, \alpha, q') > 0$, for some $q, q' \in Q\}$.

Definition 4. *[1] Let E be a (crisp) fixed set and let A be a given subset of E. An intuitionistic fuzzy set (IFS) A^+ in E is an object of the following form*

$$A^+ = \{\langle x, \mu_A(x), \nu_A(x)\rangle \big| x \in E\},$$

where the functions $\mu_A : E \to [0,1]$ and $\nu_A : E \to [0,1]$ define the value of membership and the value of non-membership of the element $x \in E$ to the set A, respectively and for every $x \in E$, $0 \leq \mu_A(x) + \nu_A(x) \leq 1$.

Obviously, every ordinary fuzzy set $\{(x, \mu_A(x)) | x \in E\}$ has an intuitionistic form $\{< x, \mu_A(x), 1 - \mu_A(x) > \big| x \in E\}$. If $\pi_A(x) = 1 - \mu_A(x) - \nu_A(x)$, then $\pi_A(x)$ is the value of non-determinacy (uncertainty) of the membership of element $x \in E$ to the set A. In the case of ordinary fuzzy sets, where $\nu_A(x) = 1 - \mu_A(x)$, we have $\pi_A(x) = 0$, for every $x \in E$.

III. Intuitionistic fuzzy multiset finite automata

Definition 5. *An intuitionistic fuzzy multiset finite automata (IFMFA) $\mathcal{M}$ is defined as: $\mathcal{M} = (Q, \Sigma, A, \iota, \tau)$, where*

1) *Q is a set of states,*
2) *Σ is a non-empty set of input alphabet,*
3) *$A = (\mu_A, \nu_A)$ is an intuitionistic fuzzy set, where $\mu_A, \nu_A : Q \times \Sigma^{\oplus} \times Q \to [0,1]$ and μ_A, ν_A are called intuitionistic fuzzy transition relation of states,*
4) *$\iota : (\iota_1, \iota_2)$ is an intuitionistic fuzzy set, $\iota_1, \iota_2 : Q \to [0,1]$ and ι_1, ι_2 are called intuitionistic initial function,*
5) *$\tau : (\tau_1, \tau_2)$ is an intuitionistic fuzzy set, $\tau_1, \tau_2 : Q \to [0,1]$ and τ_1, τ_2 are called intuitionistic output function.*

Since the operation $\oplus$ for multisets is commutative, intuitively, the instruction $\mu_A(p_1, \theta \oplus \sigma, p_2)(\nu_A(p_1, \theta \oplus \sigma, p_2))$ stands for that the current state p_1, with inputting the multiset θ or σ being scanned, can turn to the state p_2, that is, the choice of intuitionistic fuzzy multiset relation between applicable instructions (p_1, θ, p_2) and (p_1, σ, p_2) possibly being scanned different parts of the inputting multiset is non-deterministic. We next show the operation of the intuitionistic fuzzy transition relation by using the notion of a configuration. A configuration of an IFMFA is described by $(p, \theta) \in (Q, \Sigma^{\oplus})$. The IFMFA is non-deterministic, so there may be several transition relations that are possible in a given configuration. Thus, for an IFMFA $\mathcal{M}$ and its two configurations (p_1, θ) and (p_2, σ), we define a move from (p_1, θ) to (p_2, σ) with degree of membership and nonmembership $u, v \in [0,1]$, write as $(p_1, \theta) \xrightarrow{(u,v)} (p_2, \sigma)$, if there exists a multiset $\omega \in \Sigma^{\oplus}$ with $\omega \subseteq \theta$ such that $\mu_A(p_1, \omega, p_2) = u$ and $\nu_A(p_1, \omega, p_2) = v$. We use $\to^*$ to denote reflexive and transitive closure of $\to$. Let $(p_1, \theta), (p_2, \sigma) \in (Q, \Sigma^{\oplus})$, we have $(p_1, \theta) \xrightarrow{(l,k)}{}^{*} (p_2, \sigma)$ if there exists $n + 1 (n \geq 1)$ configurations $(p_1, \theta), (q_1, \theta_1), ..., (q_{n-1}, \theta_{n-1}), (p_2, \sigma)$, such that $(p_1, \theta) \xrightarrow{(l_1,k_1)} (q_1, \theta_1) \xrightarrow{(l_2,k_2)} ... \xrightarrow{(l_{n-1},k_{n-1})} (q_{n-1}, \theta_{n-1}) \xrightarrow{(l_n,k_n)} (p_2, \sigma), l_i, k_i \in [0,1]$, then $l = l_1 \wedge l_2 \wedge ... \wedge l_{n-1} \vee l_n$ and $k = k_1 \vee k_2 \vee ... \vee k_n$. Naturally the membership degrees and nonmembership degrees of a configuration (p_1, θ) deriving another configuration (p_2, σ) are expressed as follows:

$$\mu_M((p_1,\theta) \to^* (p_2,\sigma)) = \vee\{\mu_{\mathcal{M}}((p_1,\theta)) \to^* (s, \theta \ominus \omega)) \wedge \mu_{\mathcal{M}}((s, \theta \ominus \omega)) \to^* (p_2,\sigma)) \big| s \in Q, \omega \in \Sigma^{\oplus}, \omega \subseteq \theta\},$$

$$\nu_M((p_1,\theta) \to^* (p_2,\sigma)) = \wedge\{\nu_{\mathcal{M}}((p_1,\theta)) \to^* (s, \theta \ominus \omega)) \vee \nu_{\mathcal{M}}((s, \theta \ominus \omega)) \to^* (p_2,\sigma)) \big| s \in Q, \omega \in \Sigma^{\oplus}, \omega \subseteq \theta\},$$

and

$$\mu_{\mathcal{M}}((p_1,\theta)) \to^* (p_2,\theta)) = \begin{cases} 1 & \text{if } p_1 = p_2 \\ 0 & \text{if } p_1 \neq p_2 \end{cases}, \quad (2)$$

$$\nu_{\mathcal{M}}((p_1,\theta)) \to^* (p_2,\theta)) = \begin{cases} 0 & \text{if } p_1 = p_2 \\ 1 & \text{if } p_1 \neq p_2 \end{cases}. \quad (3)$$

Definition 6. *Let $\mathcal{M} = (Q, \Sigma, A, \iota, \tau)$ be an IFMFA. Let $p, q \in Q$. Then p is called an immediate successor of q with threshold (a, b) if there exists $\alpha \in \Sigma$ such that $\mu_{\mathcal{M}}((q, \alpha) \to (p, 0_\Sigma)) \geq a$ and $\nu_{\mathcal{M}}((q, \alpha) \to (p, 0_\Sigma)) \leq b$. Also, we say that p is a successor of q with threshold (a, b) if there exists $\beta \in \Sigma^{\oplus}$ such that $\mu_{\mathcal{M}}((q, \beta) \to (p, 0_\Sigma)) \geq a$ and $\nu_{\mathcal{M}}((q, \beta) \to (p, 0_\Sigma)) \leq b$, where $a, b \in [0,1]$ and $0 \leq a + b \leq 1$.*

Theorem 1. *Let $\mathcal{M} = (Q, \Sigma, A, \iota, \tau)$ be a IFMFA, $p, q, r \in Q$. Then the following hold:*

1) *q is a successor of q with threshold (a, b), where $a, b \in [0, 1]$ and $0 \leq a + b \leq 1$,*
2) *if p is a successor of q with threshold (a_1, b_1) and r is a successor of p with threshold (a_2, b_2), then r is a successor of q with threshold (a, b), where $a = a_1 \wedge a_2, b = b_1 \vee b_2$.*

Definition 7. *Let $\mathcal{M} = (Q, \Sigma, A, \iota, \tau)$ be an IFMFA and let $q \in Q$. We denote by $S^{(a,b)}(q)$ the set of all successor of q with threshold (a, b).*

Definition 8. *Let $\mathcal{M} = (Q, \Sigma, A, \iota, \tau)$ be an IFMFA and $T \subseteq Q$. The set of all successors of T with threshold (a, b) is defined as follows: $S^{(a,b)}(T) = \cup_{q \in T} S^{(a,b)}(q)$, where $a, b \in [0, 1]$ and $0 \leq a + b \leq 1$.*

Theorem 2. *Let $\mathcal{M} = (Q, \Sigma, A, \iota, \tau)$ be an IFMFA and E and F be the subsets of Q. Then the following holds:*

1) *If $E \subseteq F$, then $S^{(a,b)}(E) \subseteq S^{(a,b)}(F)$,*
2) *$E \subseteq S^{(a,b)}(E)$,*
3) *$S^{(a,b)}(S^{(a,b)}(E)) = S^{(a,b)}(E)$,*
4) *$S^{(a,b)}(A \cup B) = S^{(a,b)}(A) \cup S^{(a,b)}(B)$,*
5) *$S^{(a,b)}(A \cap B) \subseteq S^{(a,b)}(A) \cap S^{(a,b)}(B)$.*

In the next example we show that

$$S^{(a,b)}(A \cap B) \neq S^{(a,b)}(A) \cap S^{(a,b)}(B).$$

Example 1. *Let $\mathcal{M} = (Q, \Sigma, A, \iota, \tau)$ be an IFMFA. Let $Q = \{p_1, p_2, p_3, p_4\}, \Sigma = \{u, v\}$ and A is defined as follows:*

$$\begin{aligned}
&A(p_1, < u >, p_1) = (0.3, 0.7),\\
&A(p_2, < u >, p_1) = (0.7, 0.1),\\
&A(p_3, < v >, p_1) = (0.8, 0.1),\\
&A(p_1, < u > \oplus < v >, p_1) = (1, 0),\\
&A(p_4, < u >, p_4) = (0.2, 0.3),\\
&A(p_4, < v >, p_4) = (0.5, 0.5).
\end{aligned}$$

Let $A = \{p_2\}$ and $B = \{p_3\}$ and $a = 0.2, b = 0.8$. Then we have $S^{(0.2,0.8)}(A) = \{p_1\}, S^{(0.2,0.8)}(B) = \{p_1\}$. Then $S^{(0.2,0.8)}(A) \cap S^{(0.2,0.8)}(B) = \{p_1\}$, but $A \cap B = \emptyset$ so $S^{(0.2,0.8)}(\emptyset) = \emptyset$. Hence, $S^{(0.2,0.8)}(A) \cap S^{(0.2,0.8)}(B) \neq S^{(0.2,0.8)}(A \cap B)$.

Definition 9. *Let $\mathcal{M} = (Q, \Sigma, A, \iota, \tau)$ be an IFMFA. Let $p, q \in Q$ and $T \subseteq Q$. We say that $\mathcal{M}$ is a swap IFMFA with threshold (a, b) if $p \in S^{(a,b)}(T \cup \{q\})$, $p \notin S^{(a,b)}(T)$, then $q \in S^{(a,b)}(T \cup \{p\})$, where $a, b \in [0, 1]$ and $0 \leq a + b \leq 1$.*

Theorem 3. *Let $\mathcal{M} = (Q, \Sigma, A, \iota, \tau)$ be an IFMFA. Then the following are equivalent:*

1) *$\mathcal{M}$ is a swap IFMFA with threshold (a, b),*
2) *$q \in S^{(a,b)}(p)$ if and only if $p \in S^{(a,b)}(q)$, for every $p, q \in Q$,*

where $a, b \in [0, 1]$ and $0 \leq a + b \leq 1$.

Definition 10. *Let $\mathcal{M} = (Q, \Sigma, A, \iota, \tau)$ be an IFMFA and $T \subseteq Q$. Let φ be a intuitionistic fuzzy subset of $T \times \Sigma \times T$, ι' be a intuitionistic fuzzy subset of T, τ' be a intuitionistic fuzzy subset of T and suppose that $\mathcal{N} = (T, \Sigma, \varphi, \iota', \tau')$. The IFMFA $\mathcal{N}$ is called an intuitionistic fuzzy multiset finite subautomata (IFMFSA) of $\mathcal{M}$ if*

1) *$\delta|_{T \times \Sigma \times T} = \varphi$,*
2) *$S^{(0,1)}(T) \subseteq T$,*
3) *$\iota|_T = \iota'$,*
4) *$\tau|_T = \tau'$.*

It is clear that, if $\mathcal{M}$ is an intuitionistic fuzzy multiset finite subautomata of $\mathcal{N}$ and $\mathcal{N}$ is an intuitionistic fuzzy multiset finite subautomata of $\mathcal{R}$, then $\mathcal{M}$ is an IFMFSA of $\mathcal{R}$.

Example 2. *Let IFMFA $\mathcal{M} = (Q, \Sigma, A, \iota, \tau)$ as Example 1. Let $T = \{p_1, p_2, p_3\}, \Sigma = \{u, v\}$. Then we have φ as follows:*

$$\begin{aligned}
&A(p_1, < u >, p_1) = (0.3, 0.7),\\
&A(p_2, < u >, p_1) = (0.7, 0.1),\\
&A(p_3, < v >, p_1) = (0.8, 0.1),\\
&A(p_1, < u > \oplus < v >, p_1) = (1, 0).
\end{aligned}$$

Theorem 4. *Let $\mathcal{M} = (Q, \Sigma, A, \iota, \tau)$ be an IFMFA. Let $\mathcal{M}_i = (Q_i, \Sigma, A_i, \iota_i, \tau_i), i \in I$ be a family of IFMFSA of $\mathcal{M}$, where $Q_i \subseteq Q$. Then the following claim hold:*

1) *$\cap_{i \in I} \mathcal{M}_i = (\cap_{i \in I} Q_i, \Sigma, \cap_{i \in I} A_i, \cap_{i \in I} \iota_i, \cap_{i \in I} \tau_i)$ is an IFMFSA of $\mathcal{M}$,*
2) *$\cup_{i \in I} \mathcal{M}_i = (\cup_{i \in I} Q_i, \Sigma, \varrho, \iota', \tau')$ is an IFMFSA of $\mathcal{M}$, where $\varrho = A|_{\cup_{i \in I} Q_i \times \Sigma \times \cup_{i \in I} Q_i}$, $\iota' = \iota|_{\cup_{i \in I} Q_i}$ and $\tau' = \tau|_{\cup_{i \in I} Q_i}$.*

Definition 11. *Let $\mathcal{M} = (Q, \Sigma, A, \iota, \tau)$ be an IFMFA. Then $\mathcal{M}$ is called strongly connected with threshold (a, b) if for every $p, q \in Q$, $p \in S^{(a,b)}(q)$, where $a, b \in [0, 1]$.*

Definition 12. *Let $\mathcal{M} = (Q, \Sigma, A, \iota, \tau)$ be an IFMFA and $\mathcal{N} = (T, \Sigma, B, \iota', \tau')$ be an IFMFSA of $\mathcal{M}$. Then we say that $\mathcal{N}$ is nontrivial if $T \neq Q$ and $T \neq \emptyset$.*

Theorem 5. *Let $\mathcal{M} = (Q, \Sigma, A, \iota, \tau)$ be an IFMFA. Then $\mathcal{M}$ is strongly connected if and only if $\mathcal{M}$ has not nontrivial IFMSA.*

IV. Conclusion

In this note, we introduced the concept of intuitionistic fuzzy multiset finite automata. Later, we presented the notion of successor with threshold (a, b), for every state q, where $a, b \in [0, 1]$ and $0 \leq a + b \leq 1$. In addition, considering the concept of successor, we have given the notion of intuitionistic fuzzy multiset finite subautomata. Moreover, we have shown that the union and intersection of intuitionistic fuzzy multiset finite subautomata is an intuitionistic fuzzy multiset finite subautomata, too.

References

[1] K. Atanassov, " Intuitionistic fuzzy sets," Fuzzy Sets and Systems, Vol. 20. pp. 87-96, 1986.

[2] F. Blanchet-Sadri, C. Morgan, " Multiset and set decipherable codes, Computers & Mathematics with Applications," Vol. 41, pp. 1257-1262, 2001.

[3] W.D. Blizard, Multiset theory, " Notre Dame Journal of Formal Logic," Vol. 30, pp. 36-66, 1989.

[4] G. Ciobanu, "Mealy multiset automata," International Journal of Foundations of Computer Science, Vol. 17, pp. 111-126, 2006.

[5] E. CsuhajVarjú, C. Mart in-Vide, V. & Mitrana, "Multiset automata," In Workshop on Membrane Computing, Springer, Berlin, Heidelberg, pp. 69-83, 2000.
[6] Y. B. Jun, " Intuitionistic fuzzy finite state machines, Journal of Applied Mathematics and Computing," Vol. 17, pp. 109-120, 2005.
[7] Y. B. Jun, " Quotient structures of intuitionistic fuzzy finite state machines," Information Sciences, Vol. 177, pp. 4977-4986, 2007.
[8] E.S. Santos, "Maximin automata", Information and Control, Vol. 13, pp. 363-377, 1968.
[9] M. Shamsizadeh, M.M. Zahedi, "A note on "Quotient structures of intuitionistic fuzzy finite state machines"," Journal of Applied Mathematics and Computing, Vol. 51, pp. 413-423, 2016.
[10] M. Shamsizadeh, M.M. Zahedi, "Bisimulation of type 2 for BL-general fuzzy automata," Soft Computing, Vol. 23, pp. 9843-9852, 2019.
[11] M. Shamsizadeh, M.M. Zahedi, "Minimal and statewise minimal intuitionistic general L-fuzzy automata," Iranian Journal of Fuzzy Systems, Vol. 13, pp. 131-152, 2016.
[12] M. Shamsizadeh, M.M. Zahedi, "On reduced fuzzy multiset finite automata," Submitted.
[13] S.P, Tiwari, G. Vinay, M.K. Dubey, "On fuzzy multiset automata," Journal of Applied Mathematics and Computing, Vol. 51, pp. 643-657, 2016
[14] W. Wee, "On generalizations of adaptive algorithms and application of the fuzzy sets concept to pattern classification," (1967).
[15] L. Zadeh, "Fuzzy sets," Information and Control, Vol. 8, pp. 338-353, 1965.
[16] X. W. Zhang, Y. M. Li, " Intuitionistic fuzzy recognizers and intuitionistic fuzzy finite automata," Soft Computing, Vol. 13, pp. 611-616, 2009.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Jaccard Pseudo-Similarity of Fuzzy Parameterized Fuzzy Soft Matrices and Its Application to Diagnosis of Parkinson's Disease

Samet Memiş[1], Serdar Enginoğlu[2], Uğur Erkan[3]

[1]*Department of Computer Engineering, Faculty of Engineering and Natural Sciences, İstanbul Rumeli University, İstanbul, Turkey*, samettmemis@gmail.com
[2]*Department of Mathematics, Faculty of Arts and Sciences, Çanakkale Onsekiz Mart University, Çanakkale, Turkey*, serdarenginoglu@gmail.com
[3]*Department of Computer Engineering, Faculty of Engineering, Karamanoğlu Mehmetbey University, Karaman, Turkey*, ugurerkan@kmu.edu.tr

***Abstract*—This paper introduces a new similarity measure of fuzzy parameterized fuzzy soft matrices (*fpfs*-matrices), i.e., Jaccard pseudo-similarity of *fpfs*-matrices. We then provide its basic properties. Afterwards, we apply it to the diagnosis of Parkinson's Disease (PD), improving a machine learning (ML) approach. Next, we compare our approach with the well-known ML approaches, such as Naïve Bayes, *k*-Nearest Neighbor (*k*NN), Support Vector Machine (SVM), Fuzzy *k*NN, Decision Trees (DT), Boosted Trees (BT), Adaptive Boosting Tree (AdaBoost), and Random Forest (RF) in terms of accuracy, specificity, and sensitivity. The results manifest that the proposed approach makes a more accurate diagnosis of PD than the others.**

***Index Terms*– Soft sets; fpfs-matrices; similarity measure; Parkinson's disease; machine learning**

I. INTRODUCTION

James Parkinson, in 1817, described Parkinson's Disease (PD), which is a neuron-degenerative and chronic sickness [1] such that many people around the world have suffered from PD. Its general symptoms are related to muscle stiffness, slowness of movement, and tremors. The most common PD symptoms are fatigue, somnolence, cadence in the gait, and freezing of gait. Besides, the patients can have malfunctions in writing and speech skills [2]. Clinician assessment of these symptoms is crucial for the diagnosis of PD. Moreover, the researchers have created some datasets by gathering the clinician assessments. Thus, the computer-aided diagnosis comes into question, and various research has been carried out to diagnose PD. To this end, various models and approaches based on machine learning (ML), image processing, and signal processing have been proposed to diagnose PD [3-5].

Recently, some mathematical tools, such as fuzzy sets [6], soft sets [7], and their hybrid versions have been put forward and applied in various fields, including medical diagnosis. However, the aforesaid studies are limited due to their application to fictitious problems. Moreover, combining soft sets and fuzzy sets, their hybrid concepts, i.e., fuzzy soft sets [8,9], fuzzy parameterized soft sets [10], fuzzy parameterized fuzzy soft sets (*fpfs*-sets) [11], and fuzzy soft matrices [12], have been introduced and applied to fictitious decision-making problems [13-16]. On the other hand, fuzzy parameterized fuzzy soft matrices (*fpfs*-matrices) [17], which can model the problems containing fuzzy parameters or alternatives (objects), have been applied to the classification problem in ML [18-22] and performance-based value assignment (PVA) problem to salt-and-pepper noise removal filters [23-27]. These applications have manifested the modelling skill and potential of *fpfs*-matrices in different fields, such as medical diagnosis. Therefore, the concept of *fpfs*-matrices is worth studying. In this

paper, we define the Jaccard-based pseudo-similarity of *fpfs*-matrices. Afterwards, we apply pseudo-similarity to the diagnosis of PD. We can highlight the significant contributions and novelties of the paper as follows:

- A new similarity measure of *fpfs*-matrices, namely Jaccard pseudo-similarity, have been proposed.
- *fpfs*-matrices have been applied to the diagnosis of PD.
- An ML approach has been proposed via Jaccard-based pseudo-similarity of *fpfs*-matrices for the diagnosis of PD.

The rest of the present paper is structured as follows: Section 2 presents the basic definitions and concepts. Section 3 offers the proposed approach for PD. Section 4 compares the proposed approach with k-Nearest Neighbor (kNN) [28], Naïve Bayes [29], Support Vector Machine (SVM) [30], Fuzzy kNN [31], Decision Trees (DT) [32], Boosted Trees (BT) [33], Adaptive Boosting Tree (AdaBoost) [34], and Random Forest (RF) [35] in terms of accuracy, specificity, and sensitivity for Parkinsons[sic] and Parkinson's Disease datasets [36]. Final Section provides some conclusive remarks and makes some suggestions for further research.

II. PRELIMINARIES

This section, firstly, provides the concept of *fpfs*-matrices [17]. Across the present paper, let E be a parameter set, $F(E)$ be the set of all fuzzy sets over E, and $\mu \in F(E)$. Here, a fuzzy set is denoted by $\left\{{}^{\mu(x)}x : x \in E\right\}$ instead of $\{(x, \mu(x)) : x \in E\}$.

Definition 1. [11] Let U be a universal set, $\mu \in F(E)$, and α be a function from μ to $F(U)$. Then, the set $\{({}^{\mu(x)}x, \alpha({}^{\mu(x)}x)) : x \in E\}$, being the graphic of α, is called a fuzzy parameterized fuzzy soft set *(fpfs*-set*)* parameterized via E over U (or briefly over U).

In the present paper, the set of all the *fpfs*-sets over U is denoted by $FPFS_E(U)$. In $FPFS_E(U)$, because the $\text{graph}(\alpha)$ and α generated each other uniquely, the notations are interchangeable. Hence, α stands for an *fpfs*-set $\text{graph}(\alpha)$ as long as it does not lead to any confusion.

Example 1. Let $E = \{x_1, x_2, x_3, x_4\}$ and $U = \{u_1, u_2, u_3, u_4, u_5\}$. Then,

$$\alpha = \{({}^{0.3}x_1, \{{}^{0.5}u_1, {}^{0.7}u_4\}), ({}^{0}x_2, \{{}^{1}u_1, {}^{0.1}u_3, {}^{0.8}u_5\}), ({}^{0.9}x_3, \{{}^{0.7}u_1, {}^{0.4}u_2, {}^{0.2}u_4\}), ({}^{1}x_4, \{{}^{0.9}u_1, {}^{0.6}u_5\})\}$$

is an *fpfs*-set over U.

Definition 2. [17] Let $\alpha \in FPFS_E(U)$. Then, $[a_{ij}]$ is called the *fpfs*-matrix of α and is defined by

$$[a_{ij}] \coloneqq \begin{bmatrix} a_{01} & a_{02} & a_{03} & \dots & a_{0n} & \dots \\ a_{11} & a_{12} & a_{13} & \dots & a_{1n} & \dots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} & \dots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots \end{bmatrix}$$

such that for $i \in \{0,1,2,\cdots\}$ and $j \in \{1,2,\cdots\}$,

$$a_{ij} \coloneqq \begin{cases} \mu(x_j), & i = 0 \\ \alpha({}^{\mu(x_j)}x_j)(u_i), & i \neq 0 \end{cases}$$

Here, if $|U| = m - 1$ and $|E| = n$, then $[a_{ij}]$ has order $m \times n$.

Across the present paper, the set of all the *fpfs*-matrices parameterized via E over U is denoted by $FPFS_E[U]$.

Example 2. The *fpfs*-matrix of the *fpfs*-set α provided in Example 1 is as follows:

$$[a_{ij}] = \begin{bmatrix} 0.3 & 0 & 0.9 & 1 \\ 0.5 & 1 & 0.7 & 0.9 \\ 0 & 0 & 0.4 & 0 \\ 0 & 0.1 & 0 & 0 \\ 0.7 & 0 & 0.2 & 0 \\ 0 & 0.8 & 0 & 0.6 \end{bmatrix}$$

Definition 3. [17] Let $[a_{ij}] \in FPFS_E[U]$. For all i and j, if $a_{ij} = \lambda$, then $[a_{ij}]$ is called λ-*fpfs*-matrix and is denoted by $[\lambda]$. Here, $[0]$ and $[1]$ are called empty *fpfs*-matrix and universal *fpfs*-matrix, respectively.

Definition 4. [17] Let $[a_{ij}], [b_{ij}] \in FPFS_E[U]$. For all i and j,

- if $a_{ij} = b_{ij}$, then $[a_{ij}]$ and $[b_{ij}]$ are called equal *fpfs*-matrices and is denoted by $[a_{ij}] = [b_{ij}]$.
- if $a_{ij} \leq b_{ij}$, then $[a_{ij}]$ is called a submatrix of $[b_{ij}]$ and is denoted by $[a_{ij}] \cong [b_{ij}]$.

Moreover, if $[a_{ij}] \cong [b_{ij}]$ and $[a_{ij}] \neq [b_{ij}]$, then $[a_{ij}]$ is called a proper submatrix of $[b_{ij}]$ and is denoted by $[a_{ij}] \tilde{\subset} [b_{ij}]$.

Secondly, the definition of pseudo-similarity of *fpfs*-matrices is presented.

Definition 5. [20] Let $s: FPFS_E[U] \times FPFS_E[U] \to \mathbb{R}$ be a mapping. Then, for all $[a_{ij}], [b_{ij}] \in FPFS_E[U]$, s is pseudo-similarity over $FPFS_E[U]$ if and only if s satisfies the properties:

i. $s([a_{ij}], [a_{ij}]) = 1$

ii. $s([a_{ij}], [b_{ij}]) = s([b_{ij}], [a_{ij}])$

iii. $0 \leq s([a_{ij}], [b_{ij}]) \leq 1$

III. PROPOSED WORK FOR PARKINSON'S DISEASE

In this section, firstly, we offer a Jaccard-based pseudo-similarity of *fpfs*-matrices and provide some of its basic properties.

Proposition 1. The mapping $s_J: FPFS_E[U] \times FPFS_E[U] \to \mathbb{R}$ defined by

$$s_J\left([a_{ij}]_{m\times n}, [b_{ij}]_{m\times n}\right) = \frac{\varepsilon + \sum_{i=1}^{m}\sum_{j=1}^{n} \min\{a_{0j}a_{ij}, b_{0j}b_{ij}\}}{\varepsilon + \sum_{i=1}^{m}\sum_{j=1}^{n} \max\{a_{0j}a_{ij}, b_{0j}b_{ij}\}}$$

is a pseudo-similarity over $FPFS_E[U]$ and is called Jaccard pseudo-similarity. Here, $\varepsilon \ll 1$ is a positive constant, e.g., $\varepsilon = 0.0001$.

Proposition 2. Let $[0]_{m\times n}, [1]_{m\times n}, [\lambda]_{m\times n} \in FPFS_E[U]$ and $\lambda \in [0,1]$. Then,

i. $s_J([0],[0]) = 1$

ii. $s_J([1],[1]) = 1$

iii. $s_J([0],[\lambda]) = \frac{\varepsilon}{\varepsilon+mn\lambda}$

Proposition 3. $[a_{ij}]_{m\times n}, [b_{ij}]_{m\times n}, [c_{ij}]_{m\times n} \in FPFS_E[U]$,

$$[a_{ij}] \tilde{\subseteq} [b_{ij}] \tilde{\subseteq} [c_{ij}] \Rightarrow \left(s_J([a_{ij}],[c_{ij}]) \le s_J([a_{ij}],[b_{ij}]) \wedge s_J([a_{ij}],[c_{ij}]) \le s_J([b_{ij}],[c_{ij}])\right)$$

Secondly, we present some fundamental notations and definitions to be required to the proposed approach. Throughout this paper, let $D \coloneqq [d_{ij}]_{m\times(n+1)}$ stand for a data matrix related to PD, and its last column contain classes of the data. Here, m and n represent the numbers of the attributes and samples in D, respectively. $(D_{test})_{m_2\times n}$, $(D_{train})_{m_1\times n}$, and $(C)_{m_1\times 1}$ denote the testing matrix, training matrix, classes of the training matrix, the derived from D, respectively, such that $m_1 + m_2 = m$. $D_{i-train}$ and D_{i-test} indicate i^{th} row of D_{train} and D_{test}, respectively. Similarly, $D_{train-j}$ and D_{test-j} describe j^{th} column of D_{train} and D_{test}, respectively. $T'_{m_2\times 1}$ represents the predicted class matrix obtained from D_{train} and D_{test}. Let $I_m \coloneqq \{1,2,\dots,m\}$ and $I_m^* \coloneqq \{0,1,\dots,m\}$.

Definition 6. Let $u, v \in \mathbb{R}^n$. Then, the function $P: \mathbb{R}^n \times \mathbb{R}^n \to [-1,1]$ defined by

$$P(u,v) \coloneqq \frac{n\sum_{i=1}^n u_i v_i - (\sum_{i=1}^n u_i)(\sum_{i=1}^n v_i)}{\sqrt{\left[n\sum_{i=1}^n u_i^2 - \left(\sum_{i=1}^n u_i\right)^2\right]\left[n\sum_{i=1}^n v_i^2 - \left(\sum_{i=1}^n v_i\right)^2\right]}}$$

is called the Pearson correlation coefficient (PCC) between u and v.

Definition 7. Let D_{train} with $m_1 \times n$ and $C_{m_1\times 1}$ be a training matrix and the class vector of D_{train}. Then, fw is called the PCC-based feature weight vector of D_{train} and is denoted by

$$fw_{1j} \coloneqq \left|P\left(D_{train-j}, C\right)\right|, \quad j \in I_n$$

Definition 8. Let $u \in \mathbb{R}^n$. Then, the vector $\hat{u} \in \mathbb{R}^n$ defined by

$$\hat{u}_i \coloneqq \begin{cases} \dfrac{u_i - \min\limits_{k\in I_n}\{u_k\}}{\max\limits_{k\in I_n}\{u_k\} - \min\limits_{k\in I_n}\{u_k\}}, & \max\limits_{k\in I_n}\{u_k\} \neq \min\limits_{k\in I_n}\{u_k\} \\ 1, & \max\limits_{k\in I_n}\{u_k\} = \min\limits_{k\in I_n}\{u_k\} \end{cases}, \quad i \in I_n$$

is called normalizing vector of u.

Definition 9. Let $D = [d_{ij}]_{m\times(n+1)}$ be a data matrix, $i \in I_m$, and $j \in I_n$. Then, the matrix $\tilde{D} = [\tilde{d}_{ij}]_{m\times n}$ defined by

$$\tilde{d}_{ij} \coloneqq \begin{cases} \dfrac{d_{ij} - \min\limits_{k\in I_m}\{d_{kj}\}}{\max\limits_{k\in I_m}\{d_{kj}\} - \min\limits_{k\in I_m}\{d_{kj}\}}, & \max\limits_{k\in I_m}\{d_{kj}\} \neq \min\limits_{k\in I_m}\{d_{kj}\} \\ 1, & \max\limits_{k\in I_m}\{d_{kj}\} = \min\limits_{k\in I_m}\{d_{kj}\} \end{cases}$$

is called column normalized matrix (feature-fuzzification matrix) of D.

Definition 10. Let $(D_{train})_{m_1\times n}$ be a training matrix obtained from $D = [d_{ij}]_{m\times(n+1)}$. Then, the matrix $\tilde{D}_{train} = [\tilde{d}_{ij-train}]_{m_1\times n}$ defined by

$$\tilde{d}_{ij-train} \coloneqq \begin{cases} \dfrac{d_{ij-train} - \min\limits_{k\in I_m}\{d_{kj}\}}{\max\limits_{k\in I_m}\{d_{kj}\} - \min\limits_{k\in I_m}\{d_{kj}\}}, & \max\limits_{k\in I_m}\{d_{kj}\} \neq \min\limits_{k\in I_m}\{d_{kj}\} \\ 1, & \max\limits_{k\in I_m}\{d_{kj}\} = \min\limits_{k\in I_m}\{d_{kj}\} \end{cases}, \quad i \in I_{m_1} \text{ and } j \in I_n$$

is called column normalized matrix (feature-fuzzification matrix) of D_{train}.

Definition 11. Let $(D_{test})_{m_2\times n}$ be a training matrix obtained from $D \coloneqq [d_{ij}]_{m\times(n+1)}$. Then, the matrix $\tilde{D}_{test} = [\tilde{d}_{ij-test}]_{m_2\times n}$ defined by

$$\tilde{d}_{ij-test} \coloneqq \begin{cases} \dfrac{d_{ij-test} - \min\limits_{k\in I_m}\{d_{kj}\}}{\max\limits_{k\in I_m}\{d_{kj}\} - \min\limits_{k\in I_m}\{d_{kj}\}}, & \max\limits_{k\in I_m}\{d_{kj}\} \neq \min\limits_{k\in I_m}\{d_{kj}\} \\ 1, & \max\limits_{k\in I_m}\{d_{kj}\} = \min\limits_{k\in I_m}\{d_{kj}\} \end{cases}, \quad i \in I_{m_2} \text{ and } j \in I_n$$

is called column normalized matrix (feature-fuzzification matrix) of D_{test}.

Finally, we propose an ML approach using Jaccard-based pseudo-similarity of *fpfs*-matrices. Its pseudocode is as follows:

Algorithm 1. Pseudocode of the Proposed Approach

Input: $(D_{train})_{m_1\times n}$, $C_{m_1\times 1}$, and $(D_{test})_{m_2\times n}$

Output: $T'_{m_2\times 1}$

1. **procedure** ParkDiagJac(D_{train}, C, D_{test})
2. Calculate fw using D_{train} and C
3. Calculate feature fuzzification of D_{train}.and D_{test}, namely $\tilde{D}_{train}$ and $\tilde{D}_{test}$
4. **for** k from 1 to m_2 **do**
5. Calculate the test *fpfs*-matrix $[a_{ij}]$ using fw and $\tilde{D}_{i-test}$
6. **for** l from 1 to m_1 **do**
7. Calculate the train *fpfs*-matrix $[b_{ij}]$ using fw and $\tilde{D}_{i-train}$
8. $sm_{l1} \leftarrow s_J([a_{ij}],[b_{ij}])$
9. **end for**
10. $w \leftarrow \underset{l\in I_{m_1}}{\operatorname{argmax}}\{sm_{l1}\}$
11. $t'_{k1} \leftarrow$ the class label of w
12. **end for**
13. **return** $T'_{m_2\times 1}$
14. **end procedure**

IV. Performance Comparison of the Proposed Approach

This section, firstly, provides two UCI datasets related to PD. Secondly, it presents performance metrics to be required in performance comparison. Finally, it compares the proposed approach with the well-known ML techniques.

A. UCI Datasets Related to PD

In this subsection, the two UCI datasets related to PD [36] are detailed. The first dataset, Parkinsons[sic], comprises a variety of biomedical voice analysis from 31 people, 23 with PD. In the tables, each column is a different voice measure, and each row belongs to one of 195 voice recordings mentioned above. According to the "status" column, the main goal of comprising the data is to differentiate people with PD from healthy ones; the "status" column is set to 1 for PD and 0 for healthy. The second one was collected from 188 patients with PD (81 women and 107 men) with ages varying from 33 to 87 (65.1±10.9) at the Department of Neurology, Istanbul University. The control group comprises 64 healthy people (41 women and 23 men) with ages varying between 41 and 82 (61.1±8.9). During the process of collecting data, the researchers set the microphone to 44.1 kHz, and under the physician's supervision, each subject phonates sustainably the vowel "a" with three repetitions. At the same time, physicians gathered voice recordings. The properties of the aforesaid datasets are presented in Table I.

TABLE I. DETAILS OF PD DATASETS IN UCI DATABASE

Dataset	# Sample	# Attribute	# Class
Parkinsons[sic]	195	22	2
Parkinson's Disease	756	755	2

stands for "the number of"

B. Performance Metrics for Binary Classification

In this subsection, the mathematical notations of the performance metrics, namely accuracy, sensitivity, and specificity [38,39], are provided to compare the ML approaches mentioned above. Let $D_{test} = \{x_1, x_2, \cdots, x_n\}$, $T = \{T_1, T_2, \cdots, T_n\}$, and $T' = \{T'_1, T'_2, \cdots, T'_n\}$ be the set of n samples to be classified, the set of ground truth classes of the samples, and the set of prediction class of the samples, respectively. Then,

$$\text{Accuracy}(T, T') \coloneqq \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Sensitivity}(T, T') \coloneqq \frac{TP}{TP + FN}$$

$$\text{Specificity}(T, T') \coloneqq \frac{TN}{TN + FP}$$

such that

$$TP \coloneqq |\{x_k : 1 \in T_k \wedge 1 \in T'_k, k \in I_n\}|$$

$$TN \coloneqq |\{x_k : 0 \in T_k \wedge 0 \in T'_k, k \in I_n\}|$$

$$FP \coloneqq |\{x_k : 0 \in T_k \wedge 1 \in T'_k, k \in I_n\}|$$

$$FN \coloneqq |\{x_k : 1 \in T_k \wedge 0 \in T'_k, k \in I_n\}|$$

Here, TP, TN, FP, and FN stands for the numbers of true positive, true negative, false positive, and false negative, respectively.

C. Performance Results

This subsection compares the proposed approach with kNN, NB, SVM, Fuzzy kNN, DT, BT, AdaBoost, and RF in diagnosis results using the accuracy, sensitivity, and specificity metrics. We obtain the performance results of the compared algorithms by utilizing MATLAB R2021b and a workstation, whose properties are I(R)Xeon(R) CPU E5-1620v4@3.5GHz and 64GB RAM. Each algorithm is trained and tested by using the 5-fold cross-validation [37].

According to the results in Table II for Parkinsons[sic] and Parkinson's disease, the proposed approach achieves accuracy, sensitivity, and specificity results above %90 while others do not. Even if the specificity results of the well-known algorithms seem to be over 90%, their sensitivity results do not exceed 90%. Moreover, most of them except NB is below 80%. In medical diagnosis, accuracy, sensitivity, and specificity are most vital. For any algorithm, these three results have to approach 100%. One of them is below 90% is not acceptable in the field of medical diagnosis. It is noticed that the proposed ParkDiagJac outperforms the others related to accuracy, sensitivity, and specificity results in Table II, as well as standard deviation results in Table III.

TABLE II. PERFORMANCE COMPARISON OF THE APPROACHES ACCORDING TO DIAGNOSTIC RESULTS (IN PERCENTAGE) IN PARKINSONS[SIC] AND PARKINSON'S DISEASE DATASETS

ML Approaches	Parkinsons[sic] Dataset			Parkinson's Disease Dataset		
	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
kNN	84.82	63.40	91.69	72.07	32.94	85.39
NB	70.01	89.70	63.34	79.21	65.84	83.77
SVM	86.37	55.20	96.46	74.60	0.00	**100.00**
FkNN	85.33	66.00	91.53	71.81	39.72	82.73
DT	85.37	71.10	90.11	80.24	63.35	85.99
BT	90.79	79.40	94.59	91.54	72.24	98.12
AdaBoost	91.40	74.50	**97.04**	91.88	73.29	98.23
RF	90.89	74.90	96.09	87.67	59.74	97.20
ParkDiagJac	**94.77**	**90.10**	96.34	**96.14**	**92.81**	97.27

TABLE III. STANDARD DEVIATION COMPARISON OF THE APPROACHES ACCORDING TO DIAGNOSTIC RESULTS IN PARKINSONS AND PARKINSON'S DISEASE DATASETS

ML Approaches	Parkinsons[sic] Dataset			Parkinson's Disease Dataset		
	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
kNN	5.87	21.05	6.78	4.50	11.13	4.78
NB	9.34	16.46	13.23	4.90	11.03	6.05
SVM	6.23	23.75	4.63	0.49	0.00	0.00
Fuzzy kNN	6.31	19.25	6.70	4.63	9.82	5.39
DT	9.66	22.30	10.96	3.74	9.71	4.39
BT	6.49	18.59	6.77	2.90	11.78	1.73
AdaBoost	6.79	26.64	4.32	2.64	10.49	1.85
RF	6.03	20.59	5.62	2.92	12.06	2.02
ParkDiagJac	5.41	13.23	5.80	2.13	6.52	2.07

V. CONCLUSION

This paper, firstly, proposed a Jaccard-based pseudo-similarity of *fpfs*-matrices and provided its basic properties. Afterwards, it applied this pseudo-similarity to the medical diagnosis of PD utilizing two UCI datasets concerning PD. The results manifest that the proposed ParkDiagJac is an efficacious algorithm compared to well-known ML algorithms.

Future studies should be concentrated on improving the ParkDiagJac, and diagnosis performance approached 100% concerning accuracy, sensitivity, and specificity. It is believed that there are several ways to achieve the purpose. One of them is to utilize a new correlation coefficient to obtain more convenient weighting. The second is to define the new similarities of *fpfs*-matrices. The last is to employ more general mathematical tools than *fpfs*-matrices, i.e., intuitionistic fuzzy parameterized intuitionistic fuzzy soft matrices [40], interval-valued intuitionistic fuzzy parameterized interval-valued intuitionistic fuzzy soft sets [41], and picture fuzzy sets [42, 43].

REFERENCES

[1] J. Parkinson, "An essay on the shaking palsy," J. Neuropsychiatry Clin. Neurosci., vol. 20, no. 4, pp. 223–236, 1817.

[2] R. E. Burke, "Evaluation of the braak staging scheme for Parkinson's disease: Introduction to a panel presentation," Mov. Disord., vol. 25, no. S1, pp. 76–77, 2010.

[3] P. Drotar, J. Mekyska, I. Rektorová, L. Masarová, Z. Smékal, and M. Faundez-Zanuy, "Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease," Artif. Intell. Med., vol. 67, pp. 39–46, 2016.

[4] B. Li, C. J. L. Li, J. Zhang, and D. Meng, "Automated segmentation and reconstruction of the subthalamic nucleus in Parkinson's disease patients," Neuromodul: Technol. Neural Interface, vol. 19, no. 1, pp. 13–19, 2016.

[5] C. Camara, K. Warwick, R. Bruña, T. Aziz, F. del Pozo, and F. Maestú, "A fuzzy inference system for closed-loop deep brain stimulation in Parkinson's disease," J. Med. Syst. 39, Article Number: 155, 2015.

[6] L. A. Zadeh, "Fuzzy sets," Inf. Control., vol. 8, no. 3, pp. 338–353, 1965.

[7] D. Molodtsov, "Soft Set Theory-First Results," Comput. Math. with Appl., vol. 37, no. 4–5, pp. 19–31, 1999.

[8] P. K. Maji, R. Biswas, and A. R. Roy, "Fuzzy soft sets," J. Fuzzy Math., vol. 9, no. 3, pp. 589–602, 2001.

[9] N. Çağman, S. Enginoğlu, and F. Çıtak, "Fuzzy soft set theory and its applications," Iranian J. Fuzzy Syst., vol. 8, no. 3, pp. 137–147, 2011.

[10] N. Çağman, F. Çıtak, and S. Enginoğlu, "FP-soft set theory and its applications," Ann. Fuzzy Math. Inform., vol. 2, no. 2, pp. 219–226, 2011.

[11] N. Çağman, F. Çıtak, and S. Enginoğlu, "Fuzzy parameterized fuzzy soft set theory and its applications," Turkish J. Fuzzy Syst., vol. 1, no. 1, pp. 21–35, 2010.

[12] N. Çağman and S. Enginoğlu, "Fuzzy soft matrix theory and its application in decision making," Iranian J. Fuzzy Syst., vol. 7, no. 1, pp. 109–119, 2012.

[13] S. Enginoğlu and S. Memiş, "Comment on "Fuzzy Soft Sets" [The Journal of Fuzzy Mathematics 9(3), 2001, 14 589-602]," Int. J. Latest Eng. Res. Appl., vol. 3, no. 9, pp. 1–9, 2018.

[14] S. Enginoğlu and S. Memiş, "A configuration of some soft decision-making algorithms via *fpfs*-matrices," Cumhuriyet Sci. J., vol. 39, no. 4, pp. 871–881, 2018.

[15] S. Enginoğlu, S. Memiş, and T. Öngel, "Comment on soft set theory and uni-int decision-making [European Journal of Operational Research, (2010) 207, 848-855]," J. New Res. Sci., vol. 7, no. 3, pp. 28–43, 2018.

[16] S. Enginoğlu, S. Memiş, and B. Arslan, "Comment (2) on soft set theory and uni-int decision-making [European Journal of Operational Research, (2010) 207, 848-855]," J. New Theory, no. 25, pp. 84–102, 2018.

[17] S. Enginoğlu and N. Çağman, "Fuzzy parameterized fuzzy soft matrices and their application in decision-making," TWMS J. Appl. Eng. Math., vol. 10, no. 4, pp. 1105–1115, 2020.

[18] S. Enginoğlu and S. Memiş, "An application of fuzzy parameterized fuzzy soft matrices in data classification," 2nd International Conferences on Natural Science and Technology, pp. 68–77, 2019.

[19] S. Memiş, S. Enginoğlu, and U. Erkan, "A data classification method in machine learning based on normalised Hamming pseudo-similarity of fuzzy parameterized fuzzy soft matrices," Bilge Int. J. Sci. Technol. Res., vol. 3, pp. 1–8, December 2019.

[20] S. Memiş, S. Enginoğlu, and U. Erkan, "Numerical data classification via distance-based similarity measures of fuzzy parameterized fuzzy soft matrices," IEEE Access, vol. 9, pp. 88583–88601, 2021.

[21] S. Memiş, B. Arslan, T. Aydın, S. Enginoğlu, and Ç. Camcı, "A classification method based on Hamming pseudo-similarity of intuitionistic fuzzy parameterized intuitionistic fuzzy soft matrices," J. New Res. Sci., vol. 10 no. 2, pp. 59–76, 2021.

[22] S. Memiş, S. Enginoğlu, and U. Erkan, "A classification method in machine learning based on soft decision-making via fuzzy parameterized fuzzy soft matrices," Soft Comput., vol. 26 no. 3, pp. 1165–1180, 2022.

[23] S. Enginoğlu, S. Memiş, and N. Çağman, "A generalisation of fuzzy soft max-min decision-making method and its application to a performance-based value assignment in image denoising," El-Cezerî J. Sci. Eng., vol. 6, no. 3, pp. 466–481, 2019.

[24] S. Enginoğlu, S. Memiş, and F. Karaaslan, "A new approach to group decision-making method based on TOPSIS under fuzzy soft environment," J. New Res. Sci., vol. 8 no. 2, pp. 42–52, 2019.

[25] S. Enginoğlu and S. Memiş, "A new approach to the criteria-weighted fuzzy soft max-min decision-making method and its application to a performance-based value assignment problem," J. New Res. Sci., vol. 9, no. 1, pp. 19–36, 2020.

[26] S. Enginoğlu, T. Aydın, S. Memiş, and B. Arslan, "SDM methods' configurations (2017-2019) and their application to a performance-based value assignment problem: A follow up study," Ann. Optim. Theory Pract., vol. 4, no. 1, pp. 41–85, 2021.

[27] B. Arslan, T. Aydın, S. Memiş, and S. Enginoğlu, "Generalisations of SDM methods in *fpfs*-matrices space to render them operable in *ifpifs*-matrices space and their application to performance ranking of the noise-removal filters," J. New Theory, no. 30, pp. 88–116, 2021.

[28] T. M. Cover and P. E. Hart, "Nearest neighbor pattern classification," IEEE Trans. Inf. Theory, vol. 13, pp. 21–27, 1967.

[29] H. Trevor, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd ed., vol. Springer Series in Statistics, Springer: New York, 2009.

[30] C. Cortes, and V. Vapnik, "Support-vector networks," Mach. Learn., vol. 20, no. 3, pp. 273–297, 1995.

[31] J. M. Keller, M. R. Gray, and J. A. Givens, "A fuzzy k-nearest neighbor algorithm," IEEE Trans. Fuzzy Syst., vol. 15, pp. 580–585, 1985.

[32] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, Classification and Regression Trees, 3rd ed., CRC Press: Wadsworth, 1998.

[33] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," Ann. Stat., vol. 29, no. 5, pp. 1189–1232, 2001.

[34] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of On-line learning and an application to boosting," J. Comput. Syst. Sci., vol. 55, no. 1, pp. 119–139, 1997.

[35] L. Breiman, "Random forests," Mach. Learn., vol. 45, pp. 5–32, 2001.

[36] D. Dua, and C. Graff, UCI Machine Learning Repository, 2019. https://archive.ics.uci.edu/ml/

[37] M. Stone, "Cross-validatory choice and assessment of statistical predictions," J. R. Stat. Soc. Series B Stat. Methodol., vol. 36, pp. 111–147, 1974.

[38] T. Fawcett, "An introduction to ROC analysis," Pattern Recognit. Lett., vol. 27, pp. 861–874, 2006.

[39] U. Erkan, "A precise and stable machine learning algorithm: Eigenvalue classification (EigenClass)," Neural. Comput. Appl., vol. 33, no. 10, pp. 5381–5392, 2021.

[40] S. Enginoğlu and B. Arslan, "Intuitionistic fuzzy parameterized intuitionistic fuzzy soft matrices and their application in decision-making," Comput. Appl. Math., vol. 39, Article Number: 325, 2020.

[41] T. Aydın and S. Enginoğlu, "Interval-valued intuitionistic fuzzy parameterized interval-valued intuitionistic fuzzy soft sets and their application in decision-making," J. Ambient Intell. Humaniz. Comput., vol. 12, no. 1, pp. 1541–1558, 2021.

[42] B. C. Cuong, "Picture fuzzy sets," J. Comput. Sci. Cybern., vol. 30, no. 4, pp. 409–420, 2014.

[43] S. Memiş, "A study on picture fuzzy sets," 7th IFS and Contemporary Mathematics Conference (IFSCOM2021), pp. 125–132, 2021.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Lattice-Valued fuzzy hyper ideals in hyper residuated lattices

1st Mahmood Bakhshi, 2nd Mohadese Nazifi

1st Department of Mathematics, *University of Bojnord*, Bojnord, Iran, bakhshi@ub.ac.ir
2nd Department of Mathematics, *Golestan University*, Gorgan, Iran, mohadesenazifi1866@gmail.com

***Abstract*—In this paper, lattice-valued (strong) hyper RL-ideals in a hyper residuated lattice are introduced and several properties are given and some characterizations of them are obtained.**

***Index Terms*—Hyper residuated lattice, lattice-valued hyper ideal, lattice-valued strong hyper ideal**

I. Introduction and Preliminaries

Residuated lattices were introduced by Ward and Dilworth [12] as a bounded lattice L endowed with a residuated operation '$\star$' with the residual '$\rightarrow$' such that $(L;\star)$ forms a commutative monoid and also satisfies the adjoint property

$$a \star b \leq c \;\Leftrightarrow\; a \leq b \rightarrow c$$

The main examples of residuated lattices are MV-algebras [4] and BL-algebras [8].

The hyperstructure theory was introduced by Marty [9] at the 8th Congress of Scandinavian Mathematicians. In this definition, a hyperoperation on a set A is a function from A into the set of nonempty subsets of A. For more information about hyperstructures and applications see [5]. After that many researchers have applied hyperstructure theory to algebraic structures. It was Borzooei and his co-authors that applied the hyperstructures to algebraic logics and introduced hyper K-algebras [3]. After that Ghorbani et al. have applied the hyperstructures to MV-algebras and introduced hyper MV-algebras [6]. Mittas et al. [10] introduced the notion of a hyperlattice and superlattice. A superlattice is a partially ordered set $\langle A;\leq\rangle$ endowed with two binary hyperoperations $\vee$ and $\wedge$ satisfying the following properties: for all $a,b,c \in A$,

- $a \in (a \vee a) \cap (a \wedge a)$,
- $a \vee b = b \vee a$, $a \wedge b = b \wedge a$,
- $a \vee (b \vee c) = (a \vee b) \vee c$, $a \wedge (b \wedge c) = (a \wedge b) \wedge c$,
- $a \in (a \vee (a \wedge b)) \cap (a \wedge (a \vee b))$,
- $a \leq b$ implies $b \in a \vee b$ and $a \in a \wedge b$,
- if $a \in a \wedge b$ or $b \in a \vee b$, then $a \leq b$.

It must be noticed that if $(X,\leq)$ is a partially ordered set, the ordering $\leq$ can be extended to a binary relation $\ll$ in the power set of X as $A \ll B$ if and only if there exist $a \in A$ and $b \in B$ such that $a \leq b$.

In [13], Zahiri et al. applied the hyperstructure theory to residuated lattices and introduce the concept of a hyper residuated lattice as a structure $L = \langle L, \vee, \wedge, \odot, \rightarrow, 0, 1\rangle$, where

- $\langle L, \vee, \wedge, 0, 1\rangle$ is a bounded superlattice,
- $\langle L, \odot, 1\rangle$ is a commutative semihypergroup with 1 as the identity,
- the pair $(\odot, \rightarrow)$ satisfies the condition $a \odot b \ll c$ if and only if $b \ll a \rightarrow c$

Here, the symbol '$\leq$' means the ordering of the superlattice $\langle L, \vee, \wedge\rangle$. In a hyper residuated lattice L, an auxiliary hyperoperation '$\oplus$' can be defined as $x \oplus y = \neg x \rightarrow y$, where $\neg x = x \rightarrow 0$.

For more information we refer the reader to the references.

Definition 1: [1] Let I be a down-set of L; i.e.,
(HI) if $x \leq y$ and $y \in I$, then $x \in I$. I is called a

- *strong hyper RL-ideal* if $x \oplus y \subseteq I$, for all $x, y \in I$,
- *hyper RL-ideal* if $x \oplus y \ll I$, for all $x, y \in I$,

By $\mathbf{SI}(L)$ ($\mathbf{HI}(L)$) we mean the set of $\mathcal{L}$-(strong) hyper RL-ideals of L.

Lemma 1: [1] Let $I, A \subseteq L$ be nonempty.

(1) If $A \cap I \neq \emptyset$, then $A \ll I$.
(2) If I is a down-set, then $A \ll I$ implies that $A \cap I \neq \emptyset$.

Definition 2: [1] An element $a \in L$ is called a *scalar* with respect to the hyperopearation $\oplus$ (or $\oplus$-scalar) if $|a \oplus b| = |b \oplus a| = 1$, where the symbol $|.|$ denotes the cardinality of the set.

By $SC_\star(L)$, we mean the set of all scalar elements of L with respect to the hyperoperation $\star \in \{\odot, \oplus\}$.

In the sequel, in this paper, L and $\mathcal{L}$ will denote a hyper residuated lattice and a complete lattice, respectively.

II. Main results

Definition 3: An $\mathcal{L}$-set μ in L is called an $\mathcal{L}$-down set if it satisfies (LHI), where
(LHI) if $x \leq y$, then $\mu(x) \geq \mu(y)$, for all $x, y \in L$.

Definition 4: An $\mathcal{L}$-down set μ of L is called an

- $\mathcal{L}$-strong hyper RL-ideal if $\bigwedge_{a\in x\oplus y}\mu(a)\geq\mu(x)\wedge\mu(y)$,
- $\mathcal{L}$-hyper RL-ideal if $\bigvee_{a\in x\oplus y}\mu(a)\geq\mu(x)\wedge\mu(y)$.

Obviously, the $\mathcal{L}$-set $\mathbf{1}=\chi_L$ and $\mathbf{0}$, the zero $\mathcal{L}$-set of L, are $\mathcal{L}$-strong hyper RL-ideals and so are $\mathcal{L}$-hyper RL-ideals of L.

By $\mathbf{LS}(L)$ ($\mathbf{LH}(L)$), we mean the set of $\mathcal{L}$-(strong) hyper RL-ideals of L.

From the definition it is obvious that $\mathbf{LS}(L)\subseteq\mathbf{LH}(L)$.

Example 1: Let $L=\{0,a,b,c,1\}$ be a lattice whose Hasse diagram is below (Fig. 1), and let $x\wedge y$ and $x\vee y$ be the set of all lower bounds and upper bounds (respectively) of $\{x,y\}$. Define the hyperoperations $\odot$ and $\to$ as in Tables I and II. Then L is a hyper residuated lattice (see [2]). It is not difficult to verify that the $\mathcal{L}$-sets μ and η defined by $\mu(0)=s$, $\mu(c)=t$ and $\mu(a)=\mu(b)=\mu(1)=u$, and $\eta(0)=s$, $\eta(a)=\eta(b)=t$ and $\eta(c)=\eta(1)=u$, where $s,t,u\in\mathcal{L}$ with $s>t>u$ are $\mathcal{L}$-strong hyper RL-ideals of L.

Example 2: Consider the hyper residuated lattice $(L;\vee,\wedge,\odot,\to,0,1)$, where $L=\{0,a,b,1\}$ is a chain with the ordering $0<a<b<1$ and the hyperoperations $\vee$, $\wedge$, $\odot$ and $\to$ are defined as the tables III-V (see [2]). We define the $\mathcal{L}$-set μ by $\mu(0)=\mu(a)=t$, $\mu(b)=s$, $\mu(1)=u$, where $t,s,u\in\mathcal{L}$ with $t>s>u$. It is easy to verify that μ is an $\mathcal{L}$-hyper RL-ideal of L, which is not an $\mathcal{L}$-strong hyper RL-ideal becuase $\bigwedge_{x\in a\oplus a}\mu(x)=\mu(1)=u\ngeq t=\mu(a)\wedge\mu(a)$.

Theorem 1: Assume that A is a nonempty subset of L and $\mathcal{L}$-set μ is defined as

$$\mu(x)=\begin{cases}t & x\in A\\ s & x\in L\setminus A\end{cases}$$

where $t,s\in\mathcal{L}$ with $t>s$.

(1) A is a strong hyper RL-ideal of L if and only if μ is an $\mathcal{L}$-strong hyper RL-ideal of L.
(2) If A is a hyper RL-ideal, μ is an $\mathcal{L}$-hyper RL-ideal of L. Conversely, if μ is an $\mathcal{L}$-hyper RL-ideal with sup property, A is a hyper RL-ideal of L.

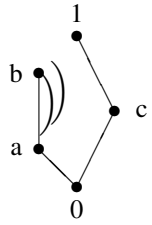


Fig. 1. The Hasse diagram of L

TABLE I
CAYLEY TABLE OF $\odot$

$\odot$	0	a	b	c	1
0	$\{0\}$	$\{0\}$	$\{0\}$	$\{0\}$	$\{0\}$
a	$\{0\}$	$\{a\}$	$\{a\}$	$\{0\}$	$\{a\}$
b	$\{0\}$	$\{a\}$	$\{a,b\}$	$\{0\}$	$\{a,b\}$
c	$\{0\}$	$\{0\}$	$\{0\}$	$\{c\}$	$\{c\}$
1	$\{0\}$	$\{a\}$	$\{a,b\}$	$\{c\}$	$\{1\}$

TABLE II
CAYLEY TABLE OF $\to$

$\to$	0	a	b	c	1
0	$\{1\}$	$\{1\}$	$\{1\}$	$\{1\}$	$\{1\}$
a	$\{c\}$	$\{1\}$	$\{1\}$	$\{c\}$	$\{1\}$
b	$\{b\}$	$\{a,b,c\}$	$\{1\}$	$\{c\}$	$\{1\}$
c	$\{a,b\}$	$\{a,b\}$	$\{a,b\}$	$\{1\}$	$\{1\}$
1	$\{0\}$	$\{1\}$	$\{a,b\}$	$\{c\}$	$\{1\}$

TABLE III
CAYLEY TABLE OF $\vee$

$\vee$	0	a	b	1
0	L	$\{a,b,1\}$	$\{b,1\}$	$\{1\}$
a	$\{a,b,1\}$	$\{a,b,1\}$	$\{b,1\}$	$\{1\}$
b	$\{b,1\}$	$\{b,1\}$	$\{b,1\}$	$\{1\}$
1	$\{0,1\}$	$\{1\}$	$\{1\}$	$\{1\}$

Proof: It is easy to verify that μ is an $\mathcal{L}$-down set if and only if A is a down-set of L. To prove sufficiency, assume that A is a (strong) hyper RL-ideal of L and $x,y\in L$. If $x\in L\setminus A$ or $y\in L\setminus A$, then $\mu(x)\wedge\mu(y)=s$. In this case, $\bigvee_{a\in x\oplus y}\mu(a)\geq\bigwedge_{a\in x\oplus y}\mu(a)\geq s=\mu(x)\wedge\mu(y)$. Now, let $x,y\in A$. In the case that A is a strong hyper RL-idea we have $x\oplus y\subseteq A$ and so for any $a\in x\oplus y$, $\mu(x)=t$, proving that μ is an $\mathcal{L}$-strong hyper RL-ideal. If A is a hyper RL-ideal, then $x\oplus y\ll A$ and so $x\oplus y\cap A\neq\emptyset$. This implies that $\bigvee_{a\in x\oplus y}\mu(a)=t\geq s=\mu(x)\wedge\mu(y)$, proving that μ is an $\mathcal{L}$-hyper RL-ideal of L.

Conversely, let $x,y\in A$. If μ is an $\mathcal{L}$-strong hyper RL-ideal of L, then $\bigwedge_{a\in x\oplus y}\mu(a)\geq\mu(x)\wedge\mu(y)=t$, whence $x\oplus y\subseteq A$, proving that A is a strong hyper RL-ideal of L. If μ is an $\mathcal{L}$-hyper RL-ideal of L, since μ satisfies the sup property, so for some $u\in x\oplus y$ we have $\mu(u)=\bigvee_{a\in x\oplus y}\mu(a)\geq\mu(x)\wedge\mu(y)=t$. This implies that $(x\oplus y)\cap A\neq\emptyset$ and so $x\oplus y\ll A$, proving that A is a hyper RL-ideal of L. ∎

Corollary 1: A nonempty subset A of L is a (strong) hyper RL-ideal of L if and only if, χ_A, the characteristic function of A, is an $\mathcal{L}$-(strong) hyper RL-ideal of L.

Proposition 1: An $\mathcal{L}$-set μ of L is an $\mathcal{L}$-(strong) hyper RL-ideal if and only if every nonempty level subset μ_t (with $t\in\mathcal{L}$) is a (strong) hyper RL-ideal of L.

Proof: Routine. ∎

Proposition 2: The intersection of any family of $\mathcal{L}$-strong hyper RL-ideals of L is again an $\mathcal{L}$-strong hyper RL-ideal.

Proof: Routine. ∎

Example 3: Consider the hyper residuated lattice L in which $L=\{x_i:i\in\mathbb{N}\}\cup\{0,1\}$ is a lattice whose Hasse diagram is below (Fig. 2) and the hyperoperations $\vee$, $\wedge$, $\odot$ and $\to$ are defined as follows (see [13]):

$$a\vee b=\{c\in L:a\leq c,b\leq c\},\ a\wedge b=\{c\in L:c\leq a,c\leq b\}$$

$a\odot b=a\wedge b$ and

$$a\to b=\begin{cases}\{1\} & : \ a\leq b\\ \{x_i:i\in\mathbb{N}\} & : \ a=1,b\in L\setminus\{1\}\\ \{x_j:j\in\mathbb{N},j\leq i\}\cup\{1\} & : \ a,b\in\{x_i:i\in\mathbb{N}\},\\ & \quad a=x_i,a\neq b\\ \{x_j:j\in\mathbb{N},j\leq i\}\cup\{1\} & : \ a\in\{x_i:i\in\mathbb{N}\},\\ & \quad b=0\end{cases}$$

TABLE IV
CAYLEY TABLE OF $\odot$

$\odot = \wedge$	0	a	b	1
0	$\{0\}$	$\{0\}$	$\{0\}$	$\{0\}$
a	$\{0\}$	$\{0,a\}$	$\{0,a\}$	$\{0,a\}$
b	$\{0\}$	$\{0,a\}$	$\{0,a,b\}$	$\{0,a,b\}$
1	$\{0\}$	$\{0,a\}$	$\{0,a,b\}$	L

TABLE V
CAYLEY TABLE OF $\rightarrow$

$\rightarrow$	0	a	b	1
0	$\{1\}$	$\{1\}$	$\{1\}$	$\{1\}$
a	$\{a,b,1\}$	$\{a,1\}$	$\{1\}$	$\{1\}$
b	$\{a,1\}$	$\{a\}$	$\{b,1\}$	$\{1\}$
1	$\{0,1\}$	$\{a\}$	$\{b,1\}$	$\{1\}$

We define $\mathcal{L}$-sets μ_j and ν_k by

$$\mu_j(x) = \begin{cases} s & x = 0 \\ t_j & x = x_j, \\ u & x \in \{x_i : i \in \mathbb{N} \setminus \{j\}\} \cup \{1\} \end{cases}$$

where $s, t_j, u \in \mathcal{L}$ with $s > t_j > u$.

It is easy to verify that for all $j \in \mathbb{N}$, μ_j is an $\mathcal{L}$-hyper RL-ideal of L. Moreover, for $j, k \in \mathcal{L}$ we have $\bigwedge_{a\in 0\oplus 0}(\mu_j \cap \mu_k)(a) = \inf_{i\in\mathbb{N}}(\mu_j \cap \mu_k)(x_i) = u \not\geq s = \mu_j(0) \wedge \mu_k(0)$. So, we conclude that the intersection of $\mathcal{L}$-hyper RL-ideals may not be an $\mathcal{L}$-hyper RL-ideal.

Remark 1: It must be noticed that although the intersection of any family of L-fuzzy strong hyper RL-ideals is again an L-fuzzy strong hyper RL-ideal, the union of them may not be an L-fuzzy strong hyper RL-ideal. For instance, consider Example 1. Then

$$\bigwedge_{w\in a\oplus c} (\mu\cup\eta)(w) = (\mu\cup\eta)(1) = u \not\geq t = (\mu\cup\eta)(a)\wedge(\mu\cup\eta)(c)$$

Let μ be an $\mathcal{L}$-set of L. By Proposition 2, the least $\mathcal{L}$-strong hyper RL-ideal of L containing μ exists, which is called the $\mathcal{L}$-strong hyper RL-ideal generated by μ denoted by $\langle\mu\rangle_s$. We observe that $\langle\mathbf{0}\rangle_s = \chi_L$. Furthermore, $\langle\cdot\rangle_s$ satisfies the following properties:

- $\mu \subseteq \langle\mu\rangle_s$,
- $\mu \subseteq \nu$ implies that $\langle\mu\rangle_s \subseteq \langle\nu\rangle_s$,
- $\langle\mu\rangle_s = \mu$ if and only if μ is an $\mathcal{L}$-strong hyper RL-ideal of L.

Thus, the mapping $\mu \mapsto \langle\mu\rangle_s$ is a closure operator on $\mathbf{LS}(L)$, where closed subsets of $\mathbf{LS}(L)$ are $\mathcal{L}$-strong hyper RL-ideals of L. Hence, by Theorem **??**, $\mathbf{LH}(L)_C = \mathbf{LS}(L)$ is

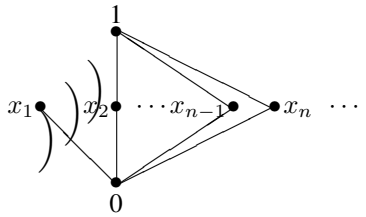


Fig. 2. Hasse diagram of L

TABLE VI
CAYLEY TABLE OF $\oplus$

$\oplus$	0	a	b	1
0	$\{0,1\}$	$\{a\}$	$\{b,1\}$	$\{1\}$
a	L	$\{a,1\}$	$\{b,1\}$	$\{1\}$
b	L	$\{a,1\}$	$\{b,1\}$	$\{1\}$
1	$\{0,1\}$	$\{a,1\}$	$\{b,1\}$	$\{1\}$

a complete lattice in which infimum and supremum is defined as

$$\bigwedge_{i\in I} \mu_i = \bigcap_{i\in I} \mu_i, \ \bigvee_{i\in I} \mu_i = \langle \cup_{i\in I}\mu_i\rangle_s. \tag{1}$$

As Example 3 shows, $\mathbf{LH}(L)$ is not closed with respect to intersection, which means that $(\mathbf{LH}(L), \subseteq)$ is not a lattice, in general. However, if the intersection is defined, similar to $\mathbf{LS}(L)$, the mapping $X \mapsto \langle X\rangle$, where $\langle X\rangle$ shows the $\mathcal{L}$-hyper RL-ideal of L generated by X, is a closure operator on $\mathbf{LS}(L)$ with $\mathbf{LH}(L)$ as the closed subsets of $\mathbf{LS}(L)$. Hence $(\mathbf{LH}(L), \subseteq)$ is a complete lattice in which supremum and infimum are defined as in (1).

Theorem 2: There exists a closure operator C on $\mathbf{LS}(L)$ such that $\mathbf{LS}(L)$ is order-isomorphic to L_C.

Proof: For $X \subseteq \mathbf{LS}(L)$ we define the mapping C as $C(X) = \{\mu \in \mathbf{LS}(L) : \mu \subseteq \bigvee X\}$. It is easy to check that C is a closure operator on $\mathbf{LS}(L)$ with $L_C = \{\mu^{\downarrow} : \mu \in \mathbf{LS}(L)\}$, where $\mu^{\downarrow} = \{\nu \in \mathbf{LS}(L) : \nu \subseteq \mu\}$. Then by [11, Theorem I.5.3], the mapping $\mu \mapsto \mu^{\downarrow}$ is the desired order-isomorphism. ∎

In the sequel, we try to give a characterization of $\mathcal{L}$-(strong) hyper RL-ideal generated by an $\mathcal{L}$-set of L.

Proposition 3: Let $\mu \in \mathbf{LH}(L)$. Then for all $x \in L$ we have

$$\begin{aligned} \mu\rangle_s(x) \quad \supseteq \quad & \bigvee\{\mu(a_1) \wedge \mu(a_2) \wedge \cdots \wedge \mu(a_n) \\ & : x \ll a_1 \oplus a_2 \oplus \cdots \oplus a_n, n \in \mathbb{N}, \\ & a_1, \ldots, a_n \in SC_{\oplus}(L)\}. \end{aligned}$$

Proof: Let $x \in L$ and $a_1, a_2, \ldots, a_n \in L$ be such that $x \ll a_1 \oplus a_2 \oplus \cdots \oplus a_n$. Then there exists $a \in a_1 \oplus a_2 \oplus \cdots \oplus a_n$ such that $x \leq a$ and hence

$$\begin{aligned} \langle\mu\rangle_s(x) \quad \geq \quad & \langle\mu\rangle_s(a) \\ = \quad & \bigcap\{\nu(a) : \nu \in \mathbf{LS}(L), \mu \subseteq \nu\} \\ \geq \quad & \bigwedge_{u\in a_1\oplus a_2\oplus\cdots\oplus a_n} \nu(u) \\ \geq \quad & \nu(a_1) \wedge \nu(a_2) \wedge \cdots \wedge \nu(a_n) \\ \geq \quad & \mu(a_1) \wedge \mu(a_2) \wedge \cdots \wedge \mu(a_n) \end{aligned}$$

Hence

$$\begin{aligned} \langle\mu\rangle_s(x) \quad \geq \quad & \bigvee\{\mu(a_1) \wedge \mu(a_2) \wedge \cdots \wedge \mu(a_n) \\ & : x \ll a_1 \oplus a_2 \oplus \cdots \oplus a_n, n \in \mathbb{N}, \\ & a_1, a_2, \ldots, a_n \in L\}. \end{aligned}$$

∎

Theorem 3: Assume that $\oplus$ is associative, $\mathbf{LH}(L)$ is closed with respect to the intersection and $\mu \in \mathbf{LH}(L)$ be such that

for all $a \in L \setminus SC_{\oplus}(L)$, $\mu(a) = 0$. Then for all $x \in L$ we have

$$\begin{aligned}\langle\mu\rangle(x) &= \bigvee\{\mu(a_1) \wedge \mu(a_2) \wedge \cdots \wedge \mu(a_n) \\ &\quad : x \ll a_1 \oplus a_2 \oplus \cdots \oplus a_n, n \in \mathbb{N}, \\ &\quad a_1, \ldots, a_n \in SC_{\oplus}(L)\}.\end{aligned}$$

Proof: Let

$$\begin{aligned}\eta(x) &= \bigvee\{\mu(a_1) \wedge \mu(a_2) \wedge \cdots \wedge \mu(a_n) \\ &\quad : x \ll a_1 \oplus a_2 \oplus \cdots \oplus a_n, n \in \mathbb{N}, \\ &\quad a_1, \ldots, a_n \in SC_{\oplus}(L)\}.\end{aligned}$$

Now, for $x \in L$, if $x \in L \setminus SC_{\oplus}(L)$, then $\mu(x) = 0 \leq \eta(x)$. Otherwise, from $x \leq x$ it is obvious that $\mu(x) \leq \eta(x)$, i.e. $\mu \subseteq \eta$. Now, let $x, y \in L$. If $x \leq y$ and $y \ll a_1 \oplus a_2 \oplus \cdots \oplus a_m$, for some $m \in \mathbb{N}$ and $a_1, a_2, \ldots, a_m \in SC_{\oplus}(L)$, so we have $x \leq y \leq a$, for some $a \in a_1 \oplus a_2 \oplus \cdots \oplus a_m$, whence $x \leq a$, by transitivity. Hence $x \ll a_1 \oplus a_2 \oplus \cdots \oplus a_m$. Thus

$$\begin{aligned}\eta(x) &= \bigvee\{\mu(a_1) \wedge \mu(a_2) \wedge \cdots \wedge \mu(a_n) : n \in \mathbb{N}, \\ &\quad a_1, a_2, \ldots, a_n \in SC_{\oplus}(L)\} \\ &\geq \mu(a_1) \wedge \mu(a_2) \wedge \cdots \wedge \mu(a_m),\end{aligned}$$

whence

$$\begin{aligned}\eta(x) &\geq \bigvee\{\mu(a_1) \wedge \mu(a_2) \wedge \cdots \wedge \mu(a_m) \\ &\quad : y \ll a_1 \oplus a_2 \oplus \cdots \oplus \mu(a_m), m \in \mathbb{N}, \\ &\quad a_1, a_2, \ldots, a_m \in SC_{\oplus}(L)\} \\ &= \eta(y).\end{aligned}$$

Now, assume that $x \ll \{u\} =_{def} c_1 \oplus c_2 \oplus \cdots \oplus c_m$ and $y \ll \{v\} =_{def} b_1 \oplus b_2 \oplus \cdots \oplus b_l$, for $m, l \in \mathbb{N}$ and $b_1, b_2, \ldots, b_l, c_1, c_2, \ldots, c_m \in SC_{\oplus}(L)$. Then

$$\begin{aligned}&x \oplus y \ll u \oplus v = \\ &((\cdots(a_1 \oplus a_2) \oplus \cdots) \oplus a_n) \oplus ((b_1 \oplus b_2) \oplus \cdots) \oplus b_m = \\ &(\cdots((((a_1 \oplus a_2) \oplus \cdots) \oplus a_n) \oplus b_1) \oplus \cdots) \oplus b_m.\end{aligned}$$

Hence, there exists $w \in x \oplus y$ such that

$$w \ll (\cdots((((a_1 \oplus a_2) \oplus \cdots) \oplus a_n) \oplus b_1) \oplus \cdots) \oplus b_m,$$

whence

$$\begin{aligned}\bigvee_{a \in x \oplus y} \eta(a) &\geq \eta(w) \\ &\geq \eta(c_1) \wedge \eta(c_2) \wedge \cdots \wedge \eta(c_m) \wedge \\ &\quad \eta(b_1) \wedge \eta(b_2) \wedge \cdots \wedge \eta(b_l) \\ &\geq \mu(a_1) \wedge \mu(a_2) \wedge \cdots \wedge \mu(a_m) \wedge \\ &\quad \mu(b_1) \wedge \mu(b_2) \wedge \mu(b_l).\end{aligned}$$

This implies that

$$\begin{aligned}\bigvee_{a \in x \oplus y} \eta(a) &\geq \bigvee\{\mu(a_1) \wedge \mu(a_2) \wedge \cdots \wedge \mu(a_m) : m \in \mathbb{N}, \\ &\quad a_1, a_2, \ldots, a_m \in SC_{\oplus}(L)\} \wedge \\ &\quad \bigvee\{\mu(b_1) \wedge \mu(b_2) \wedge \mu(b_l) : l \in \mathbb{N}, \\ &\quad b_1, b_2, \ldots, b_l \in SC_{\oplus}(L)\}.\end{aligned}$$

Now, let γ be an $\mathcal{L}$-hyper RL-ideal of L containing μ and $x \in L$ be such that $x \ll a_1 \oplus a_2 \oplus \cdots \oplus a_n$, for $n \in \mathbb{N}$ and $a_1, a_2, \ldots, a_n \in SC_{\oplus}(L)$. Then

$$\begin{aligned}\gamma(x) &\geq \gamma(a_1 \oplus a_2 \oplus \cdots \oplus a_n) \\ &\geq \gamma(a_1) \wedge \gamma(a_2) \wedge \cdots \wedge \gamma(a_n) \\ &\geq \mu(a_1) \wedge \mu(a_2) \wedge \cdots \wedge \mu(a_n).\end{aligned}$$

Hence

$$\begin{aligned}\gamma(x) &\geq \bigvee\{\mu(a_1) \wedge \mu(a_2) \wedge \cdots \wedge \mu(a_n) : n \in \mathbb{N}, \\ &\quad a_1, a_2, \ldots, a_n \in SC_{\oplus}(L)\} \\ &= \eta(x),\end{aligned}$$

means that $\eta \subseteq \gamma$. Thus η is the least $\mathcal{L}$-hyper RL-ideal of L containing μ.

■

REFERENCES

[1] M. Bakhshi, "Hyper RL-ideals in hyper residuated lattices," Discuss. Math. Gen. Algebra Appl. in press.

[2] R. A. Borzooei, M. Bakhshi, and O. Zahiri, "Filter theory of hyper residuated lattices," Quasigroups Related Systems, vol. 22, pp. 33–50, 2014.

[3] R. A. Borzooei, A. Hasankhani, M. M. Zahedi, and Y. B. Jun, "On hyper K-algebras," J. Math. Janponica, vol. 1, pp. 3305–3313, 2000.

[4] C. C. Chang, "Algebraic analysis of many-valued logics," Trans. Amer. Math. Soc. vol. 88, pp. 467–490, 1958.

[5] P. Corsini, Prolegomena of Hypergroup Theory, 2nd ed. Aviani Editore, 1993.

[6] Sh. Ghorbani, A. Hasankhani, and E. Eslami, "Hyper MV-algebra," Set-Valued Math. Appl. vol. 1, pp. 205–222, 2008.

[7] J. A. Goguen, "$\mathcal{L}$-fuzzy sets," J. Math. Anal. Appl. vol. 18, pp. 145–174, 1967.

[8] P. Hájek, Metamathematics of Fuzzy Logic, Kluwer Academic Publisher: Dordrecht, 1998.

[9] F. Marty, "Sur une generalization de la notion de groups," 8th Congress Mathematics Scandinaves, Stockholm, pp. 45–49, 1934.

[10] J. Mittas and M. Konstantinidou, "Sur un novelle génération de la notion de treillis. Les supertreillis et certaines de leurs proprités génerals," Ann. Sci. Univ. Blaise Pascal (Clermont II), Sér. Math. Fasc. vol. 25 , pp. 61–83, 1989.

[11] H. P. Sankapanavar and S. Burris, A Course in Universal Algebra, Millenuum Ed. Graduated Text Math. vol. 78, 1981.

[12] M. Ward and R. P. Dilworth, "Residuated lattices," Trans. Amer. Math. Soc. vol. 45, pp. 335–354, 1939.

[13] O. Zahiri, R. A. Borzooei, and M. Bakhshi," Quotient hyper residuated lattices," Quasigroups Related Systems, vol. 20, pp. 125–138, 2012.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Latticized-Hamacher optimization problem subject to fuzzy relational equations

Amin Ghodousian [1], Sara Zal [2]

[1] *School of Engineering Science, College of Engineering,University of Tehran, P.O.Box 11365-4563, Tehran, Iran,* a.ghodousian@ut.ac.ir
[2] *School of Engineering Science, College of Engineering, University of Tehran, P.O.Box 11365-4563, Tehran, Iran,* sara.zal@ut.ac.ir

***Abstract*— Hamacher family of t-norms is a parametric family of continuous strict t-norms, whose members are decreasing functions of the parameter. In this paper, we study a latticized optimization problem constrained by special system of fuzzy relational equations (FRE) in which fuzzy t-norms are considered as the members of the Hamacher family. First, the resolution of the feasible solutions set is investigated when it is defined with max-Hamacher composition and a necessary and sufficient condition is presented for determining the feasibility. Then, based on some theoretical properties of the problem, an algorithm is presented for solving this nonlinear problem. It is proved that the algorithm can find the exact optimal solution and an example is presented to illustrate the proposed algorithm.**

Index Terms —Fuzzy relational equations, Nonlinear optimization, Hamacher t-norm, Latticized programming.

I. INTRODUCTION

Resolution of fuzzy relational equations (FRE) with max-min composition was first studied by Sanchez [9]. Besides, Sanchez developed the application of FRE in medical diagnosis in biotechnology. Nowadays, it is well known that many issues associated with a body knowledge can be treated as FRE problems [7]. Since then the composition operator in FREs was replaced by max-product and furthermore extended to the general max-t-norm composition operator. The resolution method was kept in improving. In fact, if the max-t-norm fuzzy relational equations is consistent, then its solution set is often a non-convex set that is completely determined by a unique maximum solution and a finite number of minimal solutions [8].

Most of the existing literatures in this research field focused on the resolution of FRE and its relevant optimization problems. However, only a few research works investigated the fuzzy relational inequalities (FRI) and its relevant optimization problems [2,3,5]. For example, Guo et al. [2] studied the linear programming problem with max-min FRI constraint. Based on the concept of pseudo-minimal index, Yang [4] developed a pseudo-minimal-index algorithm to minimize a linear objective function with addition-min FRI constraint. To improve the results presented in [4], Yang et al. [6] proposed the min-max programming subject to addition-min fuzzy relational inequalities. They also studied the multi-level linear programming problem with addition-min FRI constraint [7].

In general, the genetic algorithm was applied to deal with this kind of problems. However, some fuzzy relation nonlinear optimization problems could be solved by some specific method. For example, fuzzy relation geometric programming problem was introduced by Yang and Cao [1]. Yang et al. [8] studied the single-variable term semi-latticized geometric programming subject to max-product fuzzy relation equations. The proposed problem was devised from the peer-to-peer network system and the target was to minimize the biggest dissatisfaction degrees of the terminals in such system. Yang et al. [5] introduced another version of the latticized programming problem subject to max-prod fuzzy relation inequalities with application in the optimization management model of wireless communication emission base stations.

The latticized problem was defined by minimizing objective function $z(x) = \max\{x_1, x_2, \dots, x_n\}$ subject to feasible region $X(A, b) = \{x \in [0,1]^n : \ A \circ x \geq b\}$ where "$\circ$" denotes fuzzy max-product composition.

In this paper, we study the following non-linear optimization problem in which the objective function is defined as the maximum components function and the constraints are formed as the Hamacher fuzzy relational equalities:

$$\begin{aligned} min \quad & z(x) = \max\{x_1, x_2, \dots, x_n\} \\ & A \circ x = b \\ & x \in [0,1]^n \end{aligned} \tag{1}$$

where $I = \{1,2,\dots,m\}$, $J = \{1,2,\dots,n\}$, $A = (a_{ij})_{m\times n}$, $0 \le a_{ij} \le 1$ ($\forall i \in I$ and $\forall j \in J$), is a fuzzy matrix, $b = (b_i)_{m\times 1}$, $0 \le b_i \le 1$ ($\forall i \in I$), is an m–dimensional fuzzy vector, and "$\circ$" is the max-Hamacher composition, that is, $x \circ y = H^{\alpha}(x,y) =$

$$\begin{cases} 0 & \alpha = x = y = 0 \\ \dfrac{xy}{\alpha + (1-\alpha)(x+y-xy)} & otherwise \end{cases}$$

If a_i is the i'th row of matrix A, then the constraints can be expressed as $a_i \circ x = b_i$ ($i \in I$), where $a_i \circ x = \max\limits_{j\in J}\{H^{\alpha}(a_{ij}, x_j)\}$ and

$H^{\alpha}(a_{ij}, x_j) =$

$$\begin{cases} 0 & \alpha = a_{ij} = x_j = 0 \\ \dfrac{a_{ij}\, x_j}{\alpha + (1-\alpha)(a_{ij} + x_j - a_{ij}\, x_j)} & otherwise \end{cases}$$

As mentioned, Members of the Hamacher family of t-norms are decreasing functions of the parameter α and each member of this family is actually a strict t-norm. In [2] some new operational rules of hesitant fuzzy sets were introduced based on the Hamacher t-norm and t-conorm, in which a family of hesitant fuzzy Hamacher operators was proposed for aggregating hesitant fuzzy information. In [3], the mono-tonicity of alternative scores derived from Hamacher arithmetic and geometric aggregation operators. They also investigated the relationship between alternative scores generated by Hamacher arithmetic and geometric aggregation operators. In [6], the authors focused on examining the general parametric Hamacher t-norm, where the free parameter quite essentially influences the quality of modeling and the learning capability of the model identification system.

The rest of the paper is organized as follows. In Section 2, a necessary and sufficient condition is derived to determine the feasibility of max-Hamacher FRE. In Section 3, the feasible solution set of problem (1) is characterized. It is shown that the feasible region can be expressed as the union of a finite closed convex cells. Section 4 describes the optimal solution of Problem (1). An algorithm is proposed to find the optimal solution and finally, Section 5 provides a numerical example to illustrate the algorithm.

II. FEASIBILITY CONDITION

In [6], the authors discussed some properties of FRIs, where the FRIs defined by operators with (closed) convex solutions. In this section we introduce some relevant results about the solution to system of max-Hamacher FRE. Denote $S = \{x \in [0,1]^n : A \circ x = b\}$, that is, set S represents the solution set of system (1).

Similar to the general mathematical programming problem, a vector x satisfying the constraints in (1), i.e., $x \in S$, is said to be a feasible solution of problem (1).

Definition 1. A solution $\bar{x} \in S$ is said to be the maximum solution of S when $x \le \bar{x}$ for all $x \in S$. A solution $\underline{x} \in S$ is said to be a minimal solution of S when $x \le \underline{x}$ implies $x = \underline{x}$ for any $x \in S$.

Also, for each $i \in I$, define $J_i = \{j \in J : a_{ij} \ge b_i\}$. According to [43], set S can be completely determined by one maximum solution and a finite number of minimum solutions. The maximum solution is easily obtained by $\bar{x} = \min\limits_{i\in I}\{\hat{x}_i\}$ where $\hat{x}_i = [(\hat{x}_i)_1, \dots, (\hat{x}_i)_n]$ ($\forall i \in I$) is the maximum solution of $S_i = \{x \in [0,1]^n : a_i \circ x = b_i\}$ and its components are defined as follows

$$(\hat{x}_i)_k = \begin{cases} \dfrac{[\alpha+(1-\alpha)a_{ik}]\, b_i}{a_{ik}-(1-\alpha)(1-a_{ik})\, b_i} & k \in J_i,\ b_i \ne 0 \\ 0 & k \in J_i,\ a_{ik} > b_i = 0 \\ 1 & otherwise \end{cases}$$

Moreover, if we denote the set of all minimal solutions by $\underline{S}$, then

$$S = \bigcup_{\underline{x}\in\underline{S}}\{x \in [0,1]^n : \underline{x} \le x \le \bar{x}\} \tag{2}$$

Corollary 1. $S \ne \emptyset$ if and only if $\bar{x} \in S$.

Proof. The proof is easily resulted from (2).

III. FEASIBLE REGION OF THE PROBLEM

As mentioned above, it is clear that $S = \bigcap_{i\in I} S_i$ and $\hat{x}_i$ is the maximum solution of S_i, $\forall i \in I$. The following theorem shows that the set S_i has exactly $|J_i|$ minimal solutions, where $|J_i|$ denotes the cardinality of the set J_i.

Theorem 1. Suppose that $S \ne \emptyset$, $i \in I$ and $j_0 \in J_i$. Also, define $\underline{x}(i, j_0) \in [0,1]^n$ where

$\underline{x}(i, j_0)_j =$

$$\begin{cases} \dfrac{\left[\alpha + (1-\alpha)a_{ij}\right] b_i}{a_{ij} - (1-\alpha)(1-a_{ij})\, b_i} & b_i \ne 0,\ j = j_0 \\ 0 & otherwise \end{cases}, \quad \forall j \in J$$

Then, $\underline{x}(i, j_0)$ is a minimal solution for S_i.

Proof. By contradiction, suppose that $x' \in S_i$, $x' \le \underline{x}(i, j_0)$ and $x' \ne \underline{x}(i, j_0)$. So, $x'_j \le \underline{x}(i, j_0)_j$, $\forall j \in J$ and $x' \ne \underline{x}(i, j_0)$. Therefore, $x'_j = 0$, $\forall j \in J - \{j_0\}$, and $x'_j < \frac{[\alpha+(1-\alpha)a_{ij_0}]\, b_i}{a_{ij_0}-(1-\alpha)(1-a_{ij_0})\, b_i}$. Hence, we have $H^{\alpha}(a_{ij}, x'_j) = 0$, $\forall j \in J - \{j_0\}$, and $H^{\alpha}(a_{ij_0}, x'_{j_0}) < b_i$. Therefore, $a_i \circ x' < b_i$ which contradicts $x' \in S_i$. □

Corollary 2. $S_i = \bigcup_{j\in J_i}[\underline{x}(i,j), \hat{x}_i]$, $\forall i \in I$.

Definition 2. Let $e: I \to \bigcup_{i\in I} J_i$ so that $e(i) = j \in J_i$, $\forall i \in I$, and let E be the set of all vectors e. Also, for each $e \in E$, we define $\underline{x}(e) \in [0,1]^n$ such that $\underline{x}(e)_j = \max_{i\in I} \{\underline{x}(i, e(i))_j\}$, $\forall j \in J$.

Theorem 2. $S = \bigcup_{e\in E} [\underline{x}(e), \bar{x}]$.

Proof. From Corollary 2 and the equality $S = \bigcap_{i\in I} S_i$, we have $S = \bigcap_{i\in I} \bigcup_{j\in J_i} [\underline{x}(i,j), \hat{x}_i]$, or equivalently $S = \bigcup_{e\in E} \bigcap_{i\in I} [\underline{x}(i, e(i)), \hat{x}_i]$. Therefore, $S = \bigcup_{e\in E} [\max_{i\in I} \{\underline{x}(i, e(i))\}, \min_{i\in I} \{\hat{x}_i\}]$. Now, the result follows from the definition of $\underline{x}(e)$ and $\bar{x} = \min_{i\in I} \{\hat{x}_i\}$. □

IV. Optimal solutions of the problem sing the Template

Based on Theorem 2, for each $x' \in S$ there exist $e' \in E$ such that $x' \in [\underline{x}(e'), \bar{x}]$. In other words, if $x' \le \underline{x}(e)$ and $x' \neq \underline{x}(e)$, $\forall e \in E$, then $x' \notin S$. So, we have the following corollary.

Corollary 3. $\underline{S} \subseteq \{\underline{x}(e): e \in E\}$.

The following theorem characterizes the optimal solution of problem (1).

The converse of Corollary 3 is not necessarily true (see the numerical example).

Theorem 3. If $S \neq \emptyset$,then there exists a minimal solution of S, i.e. $\underline{x}^* \in \underline{S}$, such that $\underline{x}^*$ is an optimal solution of problem (1).

Proof. Let $z(x) = \max\{x_1, x_2, \dots, x_n\}$. Furthermore, suppose that $z(\underline{x}^*) = min\{z(\underline{x}) : \underline{x} \in \underline{S}(A, b)\}$ where $\underline{x}^*$ is a minimal solution. Based on relation (3), for each $x' \in S$ there exist some $\underline{x} \in \underline{S}$ such that $\underline{x} \le x'$, i.e., $\underline{x}_j \le x'_j$, $\forall j \in J$. So, we have $max\{\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n\} \le max\{x'_1, x'_2, \dots, x'_n\}$ that implies $z(\underline{x}) \le z(x')$. But, $z(\underline{x}^*) \le z(\underline{x})$ which implies $z(\underline{x}^*) \le z(x')$, $\forall x' \in S$. □

By combination of Theorem 3 and Corollary 3, it turns out that the optimal solution of problem (1) must be a vector $\underline{x}(e^*)$ for some $e^* \in E$. Based on this fact, we can find the optimal solution of problem (1) by pairwise comparison between the elements of set $\{\underline{x}(e): e \in E\}$. We now summarize the preceding discussion as an algorithm.

Algorithm 1
Given problem (1):
1. If $\bar{x} \notin S$, then S is empty (Corollary 1).
2. Find solutions $\underline{x}(e)$,$\forall e \in E$ (Definition 2).
3. Find the minimal solutions, $\underline{S}$ by the pairwise comparison between the solutions $\underline{x}(e)$ (Corollary 3).
4. Find the optimal solution $\underline{x}(e^*)$ for problem (1) by the pairwise comparison between the objective values of the elements of $\underline{S}$ (Theorem 3).

V. Numerical example

Consider the following linear optimization problem (1):

$$min \quad z(x) = \max\{x_1, x_2, \dots, x_8\}$$

$$\begin{bmatrix} 0.0863 & 0.0521 & 0.0385 & 0.0234 & 0.0401 & 0.2026 & 0.1166 & 0.0498 \\ 0.0100 & 0.0196 & 0.0050 & 0.0289 & 0.0408 & 0.0250 & 0.6454 & 0.0214 \\ 0.6732 & 0.8699 & 0.4078 & 0.3098 & 0.7515 & 0.2230 & 0.3502 & 0.5956 \\ 0.2123 & 0.1738 & 0.3981 & 0.3237 & 0.3237 & 0.1439 & 0.2586 & 0.0747 \\ 0.0557 & 0.0230 & 0.0116 & 0.0664 & 0.0088 & 0.0439 & 0.0968 & 0.4729 \end{bmatrix} \circ \begin{bmatrix} 0.0863 \\ 0.0408 \\ 0.7515 \\ 0.3237 \\ 0.0664 \end{bmatrix}$$

$$x \in [0,1]^8$$

where "$\circ$" is the max-Hamacher composition.

Step 1: The maximum solution is obtained as
$\bar{x} = [1, 0.9061, 0.9368, 1, 1, 0.7567, 0.1403, 0.3370]$.
by a simple calculation, it is verified that $A \circ \bar{x} = b$. Therefore, the problem is feasible. On the other hand, in this example, $J_1 = \{1,6,7\}, J_2 = \{5,7\}, J_3 = \{2,5\}, J_4 = \{3,4,5\}$ and $J_5 = \{4,7,8\}$.

Step 2: According to Definition 2, $|E| = \prod_{i=1}^{5} |J_i| = 108$. Therefore, this problem has 108 minimal candidate solutions $\underline{x}(e)$.

Step 3: The problem has eight minimal solutions as follows:

$e_1 = [6, 7, 2, 4, 4]$
$\underline{x}(e_1) = [0, 0.9061, 0, 1, 0, 0.7566, 0.1403, 0]$

$e_2 = [6, 7, 2, 3, 8]$
$\underline{x}(e_2) = [0, 0.9061, 0.9368, 0, 0, 0.7566, 0.1403, 0.3367]$

$e_3 = [6, 5, 5, 4, 4]$
$\underline{x}(e_3) = [0, 0, 0, 1, 1, 0.7566, 0, 0]$

$e_4 = [6, 5, 5, 5, 8]$
$\underline{x}(e_4) = [0, 0, 0, 0, 1, 0.7566, 0, 0.3367]$

$e_5 = [1, 7, 2, 4, 4]$
$\underline{x}(e_5) = [1, 0.9061, 0, 1, 0, 0, 0.1403, 0]$

$e_6 = [1, 7, 2, 3, 8]$
$\underline{x}(e_6) = [1, 0.9061, 0.9368, 0, 0, 0, 0.1403, 0.3367]$

$e_7 = [1, 5, 5, 4, 4]$
$\underline{x}(e_7) = [1, 0, 0, 1, 1, 0, 0, 0]$

$e_8 = [1, 5, 5, 5, 8]$
$\underline{x}(e_8) = [1, 0, 0, 0, 1, 0, 0, 0.3367]$

where $e = [j_1, \dots, j_m] \in E$ means $e(i) = j_i \in J_i$, $\forall i \in I$.

Additionally, from Theorem 2 we have $S = \left[\underline{x}(e_1), \bar{x}\right] \cup \ldots \cup [\underline{x}(e_8), \bar{x}]$. So, the feasible solutions set of the problem is formed as the union of eight closed intervals.

Step 4: By a simple pairwise comparison between the objective values of the minimal solutions, we have
$0.9368 = z(\underline{x}(e_2)) < z(\underline{x}(e_i)) = 1$, for $i \in I - \{2\}$.
Therefore, $\underline{x}(e_2)$ is the unique optimal solution generated by vector e_2.

CONCLUSION

Considering the practical applications of the max-Hamacher fuzzy relational equations in FRE theory and that of the latticized programming, a nonlinear optimization problem was studied with the maximum components function as the objective function subjected to the Hamacher-FRE. Since a system of the Hamacher-FRE is a non-convex set, an algorithm was presented to find an optimal solution by using the structural properties of the problem. For this purpose, a necessary and sufficient feasibility condition was firstly derived and then, the feasible region was completely determined in terms of one maximum and a finite number of minimal solutions. It is proved that we can find the exact optimal solution of the proposed problem from the minimal solutions of the constraints, i.e., a system of max-Hamacher FRE. Additionally, a numerical example was given to illustrate the presented algorithm.

REFERENCES

[1]. Cao B. Y., The more-for-less paradox in fuzzy polynomial geometric programming, Information Sciences 211 (2012) 81-92.

[2]. Guo F. F., Pang L. P., Meng D., Xia Z. Q., An algorithm for solving optimization problems with fuzzy relational inequality constraints, Information Sciences 252 (2013) 20-31.

[3]. Guo F., Xia Z. Q., An algorithm for solving optimization problems with one linear objective function and finitely many constraints of fuzzy relation inequalities, Fuzzy Optimization and Decision Making 5 (2006) 33-47.

[4]. Li P., Fang S. C., Latticized linear optimization on the unite interval, IEEE Transactions on Fuzzy Systems 17 (2009) 1353-1365.

[5]. Li P., Wang Y., A matrix approach to latticized linear programming with fuzzy-relation inequality constraints, IEEE Transactions on Fuzzy Systems 21 (2013) 781-788.

[6]. R. Lovassy, I. J. Rudas, L. T. Kóczy, Learning the optimal parameter of the Hamacher t-norm applied for fuzzy-rule-based model extraction, Neural Computing and Applications 24 (2014),133–142.

[7]. Pedrycz W., Granular Computing: Analysis and Design of Intelligent Systems, CRC Press, Boca Raton, 2013.

[8]. Qu X. B., Wang X. P., Lei Man-hua. H., Conditions under which the solution sets of fuzzy relational equations over complete Brouwerian lattices form lattices, Fuzzy Sets and Systems 234 (2014) 34-45.

[9]. Sanchez E., Solution in composite fuzzy relation equations: application to medical diagnosis in Brouwerian logic, in: M.M. Gupta. G.N. Saridis, B.R. Games (Eds.), Fuzzy Automata and Decision Processes, North-Holland, New York, 1977, pp. 221-234.

[2]. C. Tan, W. Yi, X. Chen, Hesitant fuzzy Hamacher aggregation operators for multicriteria decision making, Applied Soft Computing (26)(2015) 325-349.

[3]. X. Tang, C. Fu, D. L. Xu, S. Yanga, Analysis of fuzzy Hamacher aggregation functions for uncertain multiple attribute decision making, Information Sciences (387)(2016) 19–33.

[4]. Yang S. J., An algorithm for minimizing a linear objective function subject to the fuzzy relation inequalities with addition-min composition, Fuzzy Sets and Systems 255 (2014) 41-51.

[5]. Yang X. P., Zhou X. G., Cao B. Y., Latticized linear programming subject to max-product fuzzy relation inequalities with application in wireless communication, Information Sciences 358–359 (2016) 44–55.

[6]. Yang X. P., Zhou X. G., Cao B. Y., Min-max programming problem subject to addition-min fuzzy relation inequalities, IEEE Transactions on Fuzzy Systems 24 (1) (2016) 111–119.

[7]. Yang X. P., Zhou X. G., Cao B. Y., Multi-level linear programming subject to addition-min fuzzy relation inequalities with application in peer-to-peer file sharing system, Journal of Intelligent and Fuzzy Systems 28 (2015) 2679–2689.

[8]. Yang X. P., Zhou X. G., Cao B. Y., Single-variable term semi-latticized fuzzy relation geometric programming with max-product operator, Information Sciences 325 (2015) 271-287.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Learning-based fuzzy c-means clustering using mixtures of Student's-t distributions with missing information

Saeed Darijani and Hamid Bigdeli

Army command and staff University, Tehran, Iran, saeed_darijani@yahoo.com

***Abstract*— In this paper is proposed a clustering algorithm, based on a fuzzy treatment of finite mixtures of multivariate Student's-t distributions, using Learning-based fuzzy c-means (LB-FCM) algorithm as well as missing data. We construct a robust LB-FCM framework for handling missing data assuming the finite mixture Student's-t distributions. Comparisons between LB-FCM and EM-type algorithms are made. Experimental results and comparisons actually demonstrate the advantage of the proposed LB-FCM.**

***Index Terms*— Asymmetry, Fuzzy clustering, Images processing, Learning-based fuzzy clustering.**

I. Introduction

Finite mixture models (FMM) are popular methods for clustering in multivariate analysis. The finite mixture of normal [1] and Student's t [2] distributions are often used in the model-based clustering analysis. The Student's t FMMs lacks robustness against missing values. This may cause misleading statistical inference results. [3] addressed missingness in the FMM using the expectation maximization (EM) algorithm introduced by [4] while [5] introduced the mixture of multivariate normal distributions for the presence of missing data. [6] considered a robust extension of FM multivariate normal models using the multivariate t distribution (FM-t) as well as incomplete data.

However, the implementation of the EM-type algorithm may also cause misleading inferential results for the model-based clustering since the conventional clustering methods restrict each point of the dataset to belong exactly to one cluster. To overcome this problem, the fuzzy clustering approach has been widely implemented. [7] proposed the fuzzy approach with the idea of uncertainty by a membership function which provides imprecise class membership information. [8] developed a new fuzzy clustering method, namely fuzzy c-means (FCM). [10] proposed robust learning of FCM by added several entropy terms in the FCM objective function. We refer to [9], [11], [12], [13] and [14] to mention a few studies in fuzzy clustering.

The aim of this paper is the development of fuzzy clustering based on mixture of multivariate Student's-t distributions [6] for handling missing values. The missing values might be outliers and this can cause seriously biased estimates and clustering and subsequently leads to distorted inference.

Motivated by these results, in this paper, we consider the proposed FCM for clustering mixture of multivariate Student's-t distributions with missing information as an alternative for heavy-tail datasets.

For parameter estimation, we use an FCM-type algorithm [10] namely Learning-based FCM (LB-FCM). Throughout, we assume that the data are missing-at-random (MAR) [15, 16], so that the missing process depends on the observed information, not the cause of the missingness. To ease the computational burden, two indicator matrices are introduced that determine the observed and missing locations of each observation separately. This methodology was follow to analyze real datasets as well as for the simulation study.

The paper is organized as follows. In section II, we review some preliminaries about the mixture of multivariate Student's-t distributions Section III outlines the parameter estimation formulation for the mixture of multivariate Student's-t distributions based on the FCM algorithm. Section IV and V respectively include experimental analysis using synthetic and real data sets to study the performance of the algorithm.

II. Finite Mixture of the Student's-t distribution with missing information

In this section, we give some preliminary results and specifically represents the finite mixture of multivariate Student's-t (FM-t) distributions.

A p-dimensional random vector $\boldsymbol{X}$ is said to follow the Student's-t distribution, denoted as $\mathrm{t}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$, where $\boldsymbol{\mu} \in \mathbb{R}^p$ is the location vector, $\boldsymbol{\Sigma}$ is the $p \times p$ positive definite the

scale covariance matrix, and $\nu \in \mathbb{R}^+$ is the degree of freedom vector, if it has the probability density function (pdf) given by

$$f_{\mathrm{t}_p}(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\Sigma},\nu) = \frac{\Gamma\left(\frac{\nu+p}{2}\right)|\boldsymbol{\Sigma}|^{1/2}}{(\pi\nu)^{p/2}\Gamma\left(\nu/2\right)\left[1+\delta(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\Sigma})/\nu\right]^{(\nu+p)/2}}, \tag{1}$$

where $\delta(x;\boldsymbol{\mu},\boldsymbol{\Sigma}) = (\boldsymbol{x}-\boldsymbol{\mu})^\top\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})$ and $\Gamma(s)$ is the Gamma function, $\Gamma(s) = \int_0^\infty e^{-z}z^{s-1}dz$. Note that if ν tends to infinity, then the pdf of $\boldsymbol{X}$ reduces to the $N_p(\boldsymbol{\mu},\boldsymbol{\Sigma})$ pdf. The Student's-t distribution can be presented by

$$\boldsymbol{X} = \boldsymbol{\mu} + \boldsymbol{X}_1/W, \quad W \perp \boldsymbol{X}_1, \tag{2}$$

where $\boldsymbol{X}_1 \sim N_p(\boldsymbol{0},\boldsymbol{\Sigma})$ and $W \sim Gamma(\nu/2,\nu/2)$ and the symbol $\perp$ indicates independence. A two-level hierarchical representation of (2) is

$$\begin{aligned} \boldsymbol{X}|W=w &\sim N_p(\boldsymbol{\mu},\boldsymbol{\Sigma}/w), \\ W &\sim Gamma(\nu/2,\nu/2). \end{aligned} \tag{3}$$

Consider a set of p-dimensional observation vectors $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n$ independently from a population with c subclasses, with the pdf of $\boldsymbol{x}_j$ given by

$$f(\boldsymbol{x}_j \mid \boldsymbol{\Theta}) = \sum_{i=1}^{c}\alpha_i f_{\mathrm{t}_p}(\boldsymbol{x}_j;\boldsymbol{\theta}_i), \tag{4}$$

where $\boldsymbol{\theta}_i = (\boldsymbol{\mu}_i,\boldsymbol{\Sigma}_i,\nu_i)$ is a vector of the unknown parameters within the i-th component ($i = 1,\ldots,c$). $(\alpha_1,\ldots,\alpha_c)$ (subjected to $\sum_{i=1}^c\alpha_i = 1$) is a vector of mixing proportions, $f_{\mathrm{t}_p}(\boldsymbol{x}_j;\boldsymbol{\theta}_i)$ is given in (1) and $\boldsymbol{\Theta} = (\alpha_1,\ldots,\alpha_{c-1},\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_c)$.

For convenience, we introduce two auxiliary permutation matrices $\boldsymbol{O}_j(p_j^o\times p)$ and $\boldsymbol{M}_j((p-p_j^o)\times p)$ to extract $\boldsymbol{X}_j^o$ and $\boldsymbol{X}_j^m$ from $\boldsymbol{X}_j$ such that $\boldsymbol{X}_j^o = \boldsymbol{O}_j\boldsymbol{X}_j$ and $\boldsymbol{X}_j^m = \boldsymbol{M}_j\boldsymbol{X}_j$. It is easy to see $\boldsymbol{O}_j^\top\boldsymbol{X}_j^o + \boldsymbol{M}_j^\top\boldsymbol{X}_j^m = \boldsymbol{X}_j$ and $\boldsymbol{O}_j^\top\boldsymbol{O}_j + \boldsymbol{M}_j^\top\boldsymbol{M}_j = \boldsymbol{I}$.

The following proposition is useful for evaluating the required conditional expectation in the FCM-type for the computational algorithm described in the next section.

Proposition 1. *From* (4) *of FM-t, we have that:*

(a) *The conditional distribution of* $\boldsymbol{X}_j^o$ *given* w_j *and* $Z_{ij} = 1$ *is*

$$\boldsymbol{X}_j^o \mid (w_j, Z_{ij}=1) \sim N_{p_j^o}(\boldsymbol{\mu}_{ij}^o, w_j^{-1}\boldsymbol{\Sigma}_{ij}^{oo}),$$

where $\boldsymbol{\mu}_{ij}^o = \boldsymbol{O}_j\boldsymbol{\mu}_i$ *and* $\boldsymbol{\Sigma}_{ij}^{oo} = \boldsymbol{O}_j\Sigma_i\boldsymbol{O}_j^\top$

(b) *The marginal distribution of the observed component* $\boldsymbol{X}_j^o$ *is*

$$f(\boldsymbol{x}_j^o \mid \boldsymbol{\Theta}) = \sum_{i=1}^{c}\alpha_i f_{t_{p_j^o}}(\boldsymbol{x}_j^o;\boldsymbol{\mu}_{ij}^o,\boldsymbol{\Sigma}_{ij}^{oo},\nu_i), \tag{5}$$

where

$$f_{t_{p_j^o}}(\boldsymbol{x}_j^o;\boldsymbol{\mu}_{ij}^o,\boldsymbol{\Sigma}_{ij}^{oo},\nu_i) = \frac{\Gamma\left(\frac{\nu_i+p_j^o}{2}\right)|\boldsymbol{\Sigma}_{ij}^{oo}|^{1/2}}{(\pi\nu_i)^{p_j^o/2}\Gamma\left(\nu_i/2\right)\left[1+\delta(\boldsymbol{x}_j^o;\boldsymbol{\mu}_{ij}^o,\boldsymbol{\Sigma}_{ij}^{oo})/\nu_i\right]^{\frac{(\nu+p_j^o)}{2}}}, \tag{6}$$

(c) *The conditional distribution of* $\boldsymbol{X}_j^m$ *given* $\boldsymbol{x}_j^o$, w_j *and* $Z_{ij} = 1$ *is*

$$\boldsymbol{X}_j^m \mid (\boldsymbol{x}_j^o, w_j, Z_{ij}=1) \sim N_{p-p_j^o}(\boldsymbol{\mu}_{ij}^{m.o}, u_j^{-1}\boldsymbol{\Sigma}_{ij}^{mm.o}),$$

where $\boldsymbol{\mu}_{ij}^{m.o} = \boldsymbol{M}_j\left(\boldsymbol{\mu}_i + \boldsymbol{\Sigma}_i\boldsymbol{S}_{ij}^{oo}\left(\boldsymbol{y}_j - \boldsymbol{\mu}_i\right)\right)$, $\boldsymbol{\Sigma}_{ij}^{mm.o} = \boldsymbol{M}_j\left(\boldsymbol{I}_p - \boldsymbol{\Sigma}_i\boldsymbol{S}_{ij}^{oo}\right)\boldsymbol{\Sigma}_i\boldsymbol{M}_j^\top$ *and* $\boldsymbol{S}_{ij}^{oo} = \boldsymbol{O}_j^\top\boldsymbol{\Sigma}_{ij}^{oo-1}\boldsymbol{O}_j$.

(d) *The conditional distribution of* W_j *given* $\boldsymbol{X}_j^o$ *and* $Z_{ij} = 1$ *is*

$$\begin{aligned} f(w_j \mid \boldsymbol{x}_j^o, Z_{ij}=1) &= \frac{w_j^{\frac{\nu_i+p_j^o}{2}-1}}{\Gamma\left(\frac{\nu_i+p_j^o}{2}\right)} \\ &\times \left(\frac{\nu_i+\delta(\boldsymbol{x}_j^o;\boldsymbol{\mu}_{ij}^o,\boldsymbol{\Sigma}_{ij}^{oo})}{2}\right)^{\frac{\nu_i+p_j^o}{2}} \\ &\times \exp\left\{-\frac{w_j}{2}\left(\nu_i+\delta(\boldsymbol{x}_j^o;\boldsymbol{\mu}_{ij}^o,\boldsymbol{\Sigma}_{ij}^{oo})\right)\right\}. \end{aligned} \tag{7}$$

Furthermore,

$$E(W_j \mid (\boldsymbol{x}_j^o, Z_{ij}=1)) = \frac{\nu_i+p_j^o}{\nu_i+\delta(\boldsymbol{x}_j^o;\boldsymbol{\mu}_{ij}^o,\boldsymbol{\Sigma}_{ij}^{oo})}, \tag{8}$$

$$\begin{aligned} E(\log(W_j) \mid (\boldsymbol{x}_j^o, Z_{ij}=1)) &= DG\left(\frac{\nu_i+p_j^o}{2}\right) \\ &- \log\left(\frac{\nu_i+\delta(\boldsymbol{x}_j^o;\boldsymbol{\mu}_{ij}^o,\boldsymbol{\Sigma}_{ij}^{oo})}{2}\right), \end{aligned} \tag{9}$$

where $DG(w) = \frac{d}{dw}\log\Gamma(w)$.

III. Proposed Clustering Algorithm

In this section, we extend the FCM-type algorithm for FM-t, to accommodate missing values in the observed data set.

A. Learning-based FCM algorithm

Following [10], we define the objective function for FM-t with complete data as

$$\begin{aligned} \boldsymbol{J}(\boldsymbol{U},\boldsymbol{\Theta}) &= \sum_{i=1}^{c}\sum_{j=1}^{n}u_{ij}\left(\frac{p}{2}\log 2\pi + \frac{1}{2}\log|\boldsymbol{\Sigma}_i^{-1}| + \frac{p-1}{2}\hat{\kappa}_{ij}^{(k)}\right. \\ &\left. + \frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}_i^{-1}\hat{\boldsymbol{\Psi}}_{ij}^{(k)}\right) - \frac{\nu_i}{2}\log\frac{\nu_i}{2} + \log\Gamma\left(\frac{\nu_i}{2}\right) - \frac{\nu_i}{2}(\hat{\kappa}_{ij}^{(k)} - \hat{w}_{ij}^{(k)})\right) \\ &- r_1\sum_{i=1}^{c}\sum_{j=1}^{n}u_{ij}\log\alpha_i + r_2\sum_{i=1}^{c}\sum_{j=1}^{n}u_{ij}\log u_{ij} \\ &- r_3 n\sum_{i=1}^{c}\alpha_i\log\alpha_i, \end{aligned} \tag{10}$$

where $r_1,\ r_2,\ r_3 > 0$,

$$\hat{\boldsymbol{\Psi}}_{ij}^{(k)} = \hat{w}_{ij}^{(k)}(\hat{\boldsymbol{x}}_{ij}^{(k)} - \boldsymbol{\mu}_i)(\hat{\boldsymbol{x}}_{ij}^{(k)} - \boldsymbol{\mu}_i)^\top + (I_p - \hat{\boldsymbol{\Sigma}}_i^{(k)}\hat{\boldsymbol{S}}_{ij}^{oo^{(k)}})\hat{\boldsymbol{\Sigma}}_i^{(k)}, \tag{11}$$

$$\hat{\boldsymbol{x}}_{ij}^{(k)} = \hat{\boldsymbol{\mu}}_i^{(k)} + \hat{\boldsymbol{\Sigma}}_i^{(k)}\hat{\boldsymbol{S}}_{ij}^{oo^{(k)}}(\boldsymbol{x}_j - \hat{\boldsymbol{\mu}}_i^{(k)}), \tag{12}$$

$$\hat{w}_{ij}^{(k)} = E(W_j|\boldsymbol{x}_j^o, Z_{ij} = 1, \hat{\boldsymbol{\Theta}}^{(k)}) = \frac{\hat{\nu}_i^{(k)} + p_j^o}{\hat{\nu}_i^{(k)} + \delta(\boldsymbol{x}_j^o; \hat{\boldsymbol{\mu}}_{ij}^{o^{(k)}}, \hat{\boldsymbol{\Sigma}}_{ij}^{oo^{(k)}})}, \tag{13}$$

$$\hat{\kappa}_{ij} = E(\log(W_j)|\boldsymbol{x}_j^o, Z_{ij} = 1, \hat{\boldsymbol{\Theta}}^{(k)}) = \mathrm{DG}\left(\frac{\hat{\nu}_i^{(k)} + p_j^o}{2}\right) - \log\left(\frac{\hat{\nu}_i^{(k)} + \delta(\boldsymbol{x}_j^o; \hat{\boldsymbol{\mu}}_{ij}^{o^{(k)}}, \hat{\boldsymbol{\Sigma}}_{ij}^{oo^{(k)}})}{2}\right) \tag{14}$$

and u_{ij} is a fuzzy membership function such that

$$0 \le u_{ij} \le 1, \quad \forall i, j, \quad \sum_{i=1}^{c} u_{ij} = 1, \quad 0 < \sum_{j=1}^{n} u_{ij} < n. \tag{15}$$

The Lagrangian function of (10) is

$$\begin{aligned}\tilde{\boldsymbol{J}}(\boldsymbol{U}, \boldsymbol{\tau}, \boldsymbol{\Theta}) = &\sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}\left(\frac{p}{2}\log 2\pi + \frac{1}{2}\log|\boldsymbol{\Sigma}_i^{-1}|\right.\\ &+ \frac{p-1}{2}\hat{\kappa}_{ij}^{(k)} + \frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}_i^{-1}\hat{\boldsymbol{\Psi}}_{ij}^{(k)}\right) - \frac{\nu_i}{2}\log\frac{\nu_i}{2}\\ &\left.+ \log\Gamma\left(\frac{\nu_i}{2}\right) - \frac{\nu_i}{2}(\hat{\kappa}_{ij}^{(k)} - \hat{w}_{ij}^{(k)})\right)\\ &- r_1\sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}\log\alpha_i + r_2\sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}\log u_{ij}\\ &- r_3 n\sum_{i=1}^{c}\alpha_i\log\alpha_i - \sum_{j=1}^{n}\tau_i\left(\sum_{i=1}^{c} u_{ij} - 1\right)\\ &- \tau_{n+1}\left(\sum_{i=1}^{c}\alpha_i - 1\right),\end{aligned} \tag{16}$$

where $\boldsymbol{\tau} = (\tau_1, \cdots, \tau_{n+1})$ are coefficients Lagrangian. Thus, the updating equation for fuzzy membership function u_{ij} is obtained as follows:

$$\hat{u}_{ij}^{(k)} = \exp\left\{\left(\hat{d}_{ij}^{(k)} + r_1\log\hat{\alpha}_i^{(k)}\right)/r_2\right\} \Big/ \sum_{i=1}^{c}\exp\left\{\left(\hat{d}_{ij}^{(k)} + r_1\log\hat{\alpha}_i^{(k)}\right)/r_2\right\}, \tag{17}$$

where

$$\begin{aligned}\hat{d}_{ij}^{(k)} = &\left(\frac{p}{2}\log 2\pi + \frac{1}{2}\log|\boldsymbol{\Sigma}_i^{-1}| + \frac{p-1}{2}\hat{\kappa}_{ij}^{(k)}\right.\\ &+ \frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}_i^{-1}\hat{\boldsymbol{\Psi}}_{ij}^{(k)}\right) - \frac{\nu_i}{2}\log\frac{\nu_i}{2} + \log\Gamma\left(\frac{\nu_i}{2}\right)\\ &\left.- \frac{\nu_i}{2}(\hat{\kappa}_{ij}^{(k)} - \hat{w}_{ij}^{(k)})\right).\end{aligned} \tag{18}$$

Minimizing the objective function (16) over α_i and under the condition $\sum_{i=1}^{c}\alpha_i = 1$, we find the update equation for α_i as following

$$\hat{\alpha}_i^{(new)} = \frac{1}{n}\sum_{j=1}^{n}\hat{u}_{ij} + \frac{r_3}{r_1}\hat{\alpha}_i^{(old)}\left(\ln\hat{\alpha}_i^{(old)} - \sum_{i=1}^{c}\hat{\alpha}_i^{(old)}\ln\hat{\alpha}_i^{(old)}\right).$$

Furthermore, the new number of components is obtained as

$$c^{(new)} = c^{(old)} \tag{19}$$

$$- |\{\hat{\alpha}_i^{(old)} :\ \hat{\alpha}_i^{(old)} < 1/n,\ i \in \{1, \ldots, c^{(old)}\}\}|. \tag{20}$$

For each $i \in \{1, \ldots, c^{(new)}\}$, we also obtain an adjusted $\hat{\alpha}_i^{(new)}$ and $\hat{u}_{ij}$ as

$$\hat{\alpha}_i^{(new)} = \frac{\hat{\alpha}_i^{(new)}}{\hat{\alpha}_1^{(new)} + \cdots + \hat{\alpha}_{c^{(new)}}^{(new)}} \tag{21}$$

and

$$\hat{u}_{ij} = \frac{\hat{u}_{ij}}{\hat{u}_{1j} + \cdots + \hat{u}_{c^{(new)}j}}, \tag{22}$$

to ensure that $\hat{\alpha}_1^{(new)} + \cdots + \hat{\alpha}_{c^{(new)}}^{(new)} = 1$ and $\hat{u}_{1j} + \cdots + \hat{u}_{c^{(new)}j} = 1$ hold. Yang and Nataliani (2017) [10] recommended using

$$r_3 = \min\left[\frac{1}{c}\sum_{i=1}^{c}\exp\{-\eta n|\hat{\alpha}_i^{(new)} - \hat{\alpha}_i^{(old)}|\}, \frac{1 - \max_{1\le i\le c}\left(\frac{1}{n}\sum_{j=1}^{n}\hat{u}_{ij}\right)}{-\hat{\pi}_{(c)}^{(old)}E}\right], \tag{23}$$

where $E = \sum_{i=1}^{c}\hat{\alpha}_i^{(old)}\ln\hat{\alpha}_i^{(old)}$, $\hat{\alpha}_{(c)}^{(old)} = \max(\hat{\alpha}_1^{(old)}, \ldots, \hat{\alpha}_c^{(old)})$, and $\eta = \min\{1, 0.5^{\lfloor p/2-1\rfloor}\}$, in which $\lfloor a\rfloor$ represents the integer part of a, with $r_1^{(k)} = \mathrm{e}^{-k/10}$ and $r_2^{(k)} = \mathrm{e}^{-k/100}$.

Update $\boldsymbol{\mu}_i$ by minimizing (16), which gives

$$\hat{\boldsymbol{\mu}}_i^{(k+1)} = \frac{\sum_{j=1}^{n}\hat{u}_{ij}^{(k)}\hat{\boldsymbol{x}}_{ij}^{(k)}}{\sum_{j=1}^{n}\hat{u}_{ij}^{(k)}\hat{w}_{ij}^{(k)}}. \tag{24}$$

$\boldsymbol{\Sigma}_i$ can be updated by minimizing (16) with $\boldsymbol{\mu}_i$ held fixed at $\hat{\boldsymbol{\mu}}_i$. This gives

$$\hat{\boldsymbol{\Sigma}}_i^{(k+1)} = \frac{\sum_{j=1}^{n}\hat{u}_{ij}^{(k)}\hat{\boldsymbol{\Psi}}_{ij}^{(k)}}{\sum_{j=1}^{n}\hat{u}_{ij}^{(k)}}. \tag{25}$$

For parameter estimation of ν_i,we solution the following equation with respect to ν_i yields

$$\log\left(\frac{\nu_i}{2}\right) + 1 - \mathrm{DG}\left(\frac{\nu_i}{2}\right) + \frac{\sum_{j=1}^{n}\hat{u}_{ij}^{(k)}(\hat{\kappa}_{ij}^{(k)} - \hat{w}_{ij}^{(k)})}{\sum_{j=1}^{n}\hat{u}_{ij}^{(k)}}. \tag{26}$$

If the dfs are assumed to be identical, namely $\nu_1 = \ldots = \nu_c = \nu$, the above parameter estimation is suggested to switch to the following equation.

$$\hat{\nu}^{(k+1)} = \arg\max_{\nu}\sum_{j=1}^{n}\log\left[\sum_{i=1}^{c}\hat{\alpha}_i^{(k)}f_{\mathrm{t}_{p_j^o}}(\boldsymbol{x}_j^o; \hat{\boldsymbol{\mu}}_i^{(k)}, \hat{\boldsymbol{\Sigma}}_i^{(k)}, \nu)\right]$$

For initial values of the parameters, assume $c^{(0)} = n$ and $\alpha_i = 1/n$ and $\nu_i^{(0)} = 50$. For the initial values of $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}_i$, we follow a similar approach as in Yang et. al (2012) [17] and set

$$\boldsymbol{\mu}_i^{(0)} = \boldsymbol{y}_j, \quad \boldsymbol{\Sigma}_i^{(0)} = d^2_{i(\lfloor\sqrt{c^{(0)}}\rfloor)}\boldsymbol{I}_d, \tag{27}$$

where $\boldsymbol{I}_d$ is a $d \times d$ diagonal matrix and

$$D_i = \text{sort}\{d_{ij}^2 = \|\boldsymbol{x}_j - \boldsymbol{\mu}_i\|^2 : \ d_{ij}^2 > 0, \\ i \neq j, 1 \leq i, j \leq n\} = \{d^2_{i(1)}, d^2_{i(2)}, \ldots, d^2_{i(n)}\}.$$

Yang et. al (2012) [17] provide possible ways of avoiding the singularity of initial values $\boldsymbol{\Sigma}_i^{(0)}$. In the following algorithm, by inputting data and initial values, the output of estimating parameters and classification indices can be obtained. Algorithm 1 outlines the steps of the LB-FCM algorithm.

Algorithm 1: The LB-FCM algorithm

1. Let $\epsilon = 10^{-5}$ and set the initial values $c^{(0)} = n$, $\alpha_i = 1/n$, $\nu_i^{(0)} = 50$, $\boldsymbol{\mu}_i^{(0)} = \boldsymbol{x}_i$, $\boldsymbol{\Sigma}_i^{(0)}$ as in (27) and initial learning rates $r_1^{(0)} = r_2^{(0)} = r_3^{(0)} = 1$.
2. Compute $\hat{\boldsymbol{U}}^{(0)} = \{\hat{u}_{ij}^{(0)}\}$ by (17), $\hat{\boldsymbol{w}}^{(0)} = \{\hat{w}_{ij}^{(0)}\}$ and $\hat{\boldsymbol{\kappa}}^{(0)} = \{\hat{\kappa}_{ij}^{(0)}\}$ by (13) and (14) with $\boldsymbol{\Theta}^{(0)}$ in Step 1 and set the iteration number k to be 1.
3. Update $\hat{\boldsymbol{\Theta}}^{(k)}$ with $(\hat{\boldsymbol{U}}^{(k-1)}, \hat{\boldsymbol{w}}^{(k-1)}, \hat{\boldsymbol{\kappa}}^{(k-1)}, \hat{\boldsymbol{x}}^{(k-1)})$ by above equations.
4. Compute $\hat{u}_{ij}^{(k)}$ using $c^{(k-1)}$, $\hat{\alpha}_i^{(k-1)}$, $\hat{\boldsymbol{\mu}}_i^{(k-1)}$, $\hat{\boldsymbol{\Sigma}}_i^{(k-1)}$, $\hat{\nu}_i^{(k-1)}$, $r_1^{(k-1)}$, $r_2^{(k-1)}$ by (17).
5. Compute $\hat{\boldsymbol{\alpha}}^{(k)}$ with $\hat{\boldsymbol{U}}^{(k-1)}$, $\hat{\boldsymbol{\alpha}}^{(k-1)}$, $\hat{r}_1^{(k-1)}$ and $\hat{r}_3^{(k-1)}$ by (19).
6. Update $r_1^{(k)}$ and $r_2^{(k)}$ by $\mathrm{e}^{-k/10}$ and $\mathrm{e}^{-k/100}$, respectively and $\hat{r}_3^{(k)}$ with $\hat{\boldsymbol{\alpha}}^{(k)}$ and $\hat{\boldsymbol{\alpha}}^{(k-1)}$ by (23).
7. Update $c^{(k-1)}$ to $c^{(k)}$ by discarding irrelevant components whose $\hat{\alpha}_i^{(k)} < 1/n$ and obtain adjusted $\hat{\alpha}_i^{(k)}$ and $\hat{u}_{ij}^{(k)}$ by (20) and (22). If $k \geq 60$ and $c^{(k-60)} - c^{(k)} = 0$, then let $\hat{r}_3^{(k)} = 0$.
8. Update $\hat{\boldsymbol{U}}^{(k)}$ by (17), $\hat{\boldsymbol{w}}^{(k)}$ and $\hat{\boldsymbol{\kappa}}^{(k)}$ by (13) and (14) with $\hat{\boldsymbol{\Theta}}^{(k)}$.
9. Update $\hat{\boldsymbol{\Theta}}^{(k+1)}$ with $(\hat{\boldsymbol{U}}^{(k)}, \hat{\boldsymbol{w}}^{(k)}, \hat{\boldsymbol{\kappa}}^{(k)})$ by the equations (21), and (24)–(26).
10. If $\max_{1 \leq i \leq c^{(k)}} \|\boldsymbol{\mu}_i^{(k+1)} - \boldsymbol{\mu}_i^{(k)}\| < \epsilon$, then stop. Else increment k by 1 and return to Step 5.

As a by-product in the implementation of the above algorithm, the imputation of missing component $\boldsymbol{x}_j^m$ can be calculated as

$$\hat{\boldsymbol{x}}_j^m = \boldsymbol{M}_j \sum_{i=1}^{c} \hat{\alpha}_i \left(\hat{\boldsymbol{\mu}}_i + \hat{\boldsymbol{\Sigma}}_i \hat{\boldsymbol{S}}_{ij}^{oo}(\boldsymbol{x}_j - \hat{\boldsymbol{\mu}}_i)\right) \tag{28}$$

IV. Synthetic Data Analysis

In this section, we focus on two experiments including simulated data and one real data set from a learning repository. In the latent experiment, we refine the data to include missing values. In comparing the classification performance of different model-based classifiers, we adopt the adjusted Rand index (ARI) and the correct classification rate (CCR) with higher values meant for good classification results. For parameter estimation based EM-type algorithm, we use method following [18].

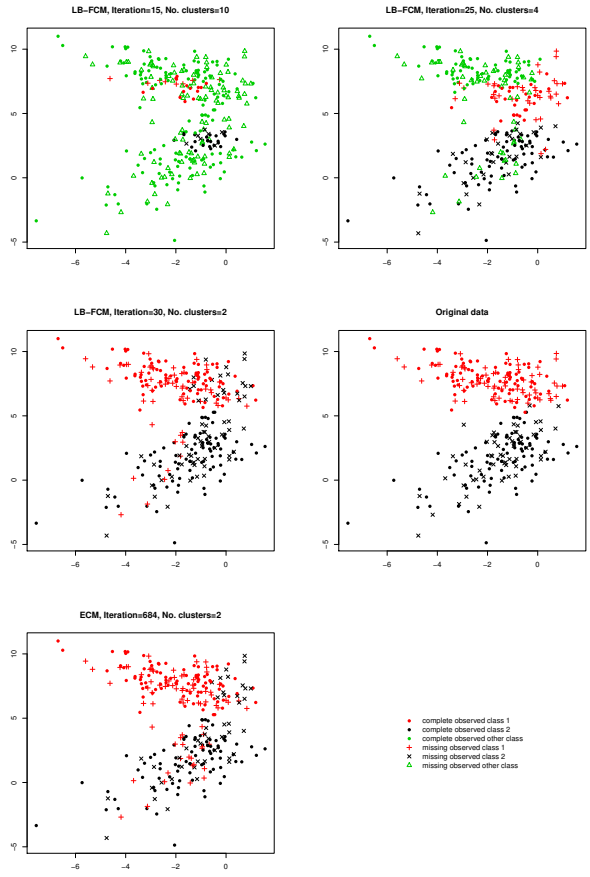


Fig. 1. Clustering results for ECM and LB-FCM algorithm (example 1).

A. Example 1

In this example, we generated data from two-component Student's t distribution with parameters $\pi_1 = \pi_2 = 1/2$, $\boldsymbol{\mu}_1 = (0, 4)^\top$, $\boldsymbol{\mu}_2 = (0, 7)^\top$, $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\nu_1 = 10$ and $\boldsymbol{\nu}_2 = 12$ with 300 data points. To proceed with the experimental study, synthetic missing values were introduced to the simulated data under the MAR mechanism. In the MAR experiment, missing items were obtained by deleting at random under $rate = 20\%$. After 15 iterations, the number of cluster decreased rapidly from 300 to 10, as shown in Fig. 1. The LB-FCM algorithm decreased the number of cluster to 4 clusters after 10 iterations. Finally, after 30 iterations (see Fig. 1), the LB-FCM algorithm obtained its converged where two clusters are formed with $c = 2$. We note that the ECM algorithm convergence after 684 iterations with two clusters. Also, the (ARI,CCR) for LB-FCM algorithm after 30 iterations and ECM algorithm were (0.587,0.884) and (0.526,0.863), respectively.

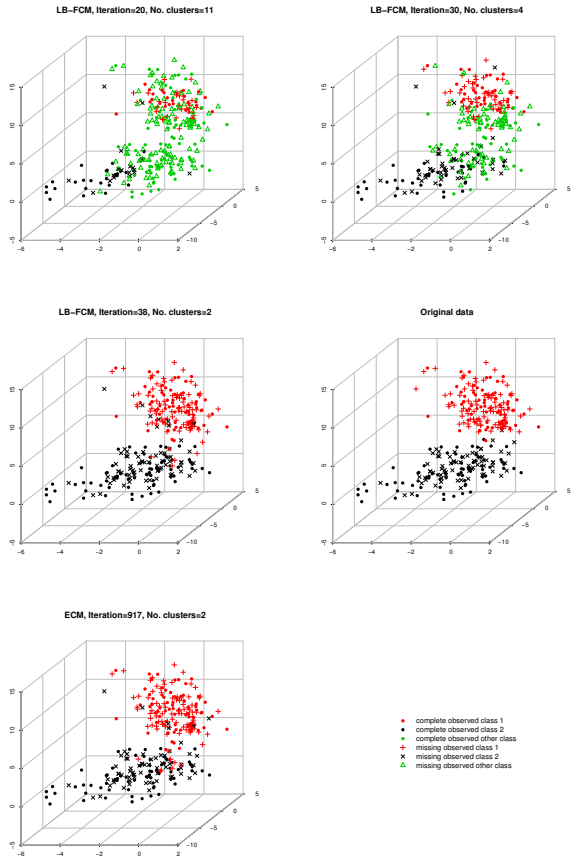


Fig. 2. Clustering results for ECM and LB-FCM algorithm (example 2).

B. Example 2

In this example, we consider a 3-dimensional data set with two blocks generated from Student's t distributions with parameters $\pi_1 = \pi_2 = \pi_3 = 1/3$, $\boldsymbol{\mu}_1 = (0,0,4)^\top$, $\boldsymbol{\mu}_2 = (0,0,6)^\top$, $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \begin{pmatrix} 1 & 0.5 & 0 \\ 0.5 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $\boldsymbol{\Sigma}_3 = \begin{pmatrix} 1 & 0.2 & 0.8 \\ 0.2 & 1 & 0.5 \\ 0.8 & 0.5 & 1 \end{pmatrix}$, $\nu_1 = 10$, $\nu_2 = 8$ and $\nu_3 = 12$, where each block contains 150 data points, as shown in Fig. 2. To proceed with the experimental study, synthetic missing values were introduced to the simulated data under the MAR mechanism. In the MAR experiment, missing items were obtained by deleting at random under high $rate = 30\%$ rates of missingness. Using the LB-FCM algorithm, 11 clusters were obtained after 20 iterations, as shown in Fig. 2. LB-FCM detected 2 clusters with ARI = 0.798 and CCR=0.947, while using ECM with $c = 2$, the ARI and CCR after 917 iterations were 0.761 and 0.936, respectively.

V. Image Reconstruction

The our method has been well recognized as a tool for perform image compression. In this section, we apply the LB-FCM method for parameter estimation and clustering

Fig. 3. Original and clustering using LB-FCM Lena image.

data based on FM-t distribution with missing values as a dramatically flexible alternative to image compression and compare the quality of the reconstructed images our model. In this example, we use Lenna image with 512×512 pixels from Matlab database with 60% missing data. This image is subdivided into $n = 4096$ non-overlapping blocks of 8×8 pixels and each block is taken as a 64-dimensional vector $\boldsymbol{x}_j$, $j = 1, \ldots, 4096$. The data vectors were then fitted under LB-FCM scenarios. Fig. 3 shows example of Lena image contaminated with 60% missing values together with eight reconstructed images obtained by fitted the LB-FCM FM-t distribution. Using the LB-FCM algorithm, we obtained seven clusters ($c = 7$) with 52 iterations, as shown in Fig. 3, but the ECM algorithm with $c = 7$ convergence after 1279 iterations.

VI. Conclusions

Our contribution enhances the existing literature since we proposed a new algorithm for clustering multivariate heavy-tailed data as well as missing data. First, we reviewed some preliminary results and specifically the representation of the finite mixture of Student's t distribution with the FCM algorithm for parameter estimation and clustering. Followed by the development of fuzzy algorithms, especially the fuzzy c-means algorithm, to obtain parameter estimates and clustering based on the mixture of Student's t distribution in the presence of missing values. We refer to it as a learning-based FCM clustering algorithm. The proposed LB-FCM uses the number of data points as initial number of clusters for solving the initialization problem. The main difference between the LB-FCM and ECM algorithm for the mixture of Student's t distributions is the beginning iteration when assigning the number of data points as initial number of clusters with missing data. Several numerical data with image segmentation are used to showcase the advantage of the LB-FCM. For our future work, we will consider data with high dimensions by building a new feature selection procedure in the LB-FCM algorithm for scale mixture of multivariate skew normal distribution.

References

[1] R. A. Redner and H. F. Walker. "Mixture densities, maximum likelihood and the em algorithm". SIAM review, vol. 26, 1984, pp. 195–239.

[2] D. Peel and G. J. McLachlan. "Robust mixture modelling using the t distribution". Stat. and comput. vol. 10, 2000, pp. 339–348.

[3] Z. Ghahramani and M. I. Jordan. "Supervised learning from incomplete data via an em approach". Advan neur info. proc. syst., 1994, pp. 120–127.

[4] A. P. Dempster and N. M. Laird and D. B. Rubin. "Maximum likelihood from incomplete data via the em algorithm". Jour. of the Roy. Stat. Soci. Seri. B (Meth.),1997 vol. 39, 1997, pp. 1–38.

[5] T. I. Lin, J. C. Lee and H. J. Ho. "On fast supervised learning for normal mixture models with missing information". Patt. Recog., vol. 39, 2006, pp. 1177–1187.

[6] H. Wang, Q. bing Zhang, B. Luo and S. Wei. "Robust mixture modelling using multivariate t-distribution with missing information". Patt. Recog., vol. 25, 2004, pp. 701–710.

[7] L. A. Zadeh. "Fuzzy sets". Info. and cont. vol. 8, 1965, pp. 338–353.

[8] D. E. Gustafson and W.C. Kessel. "Fuzzy clustering with a fuzzy covariance matrix". IEEE conf. on deci. and conto. inclu. the 17th sympo. on adap. proc., 1979, pp. 761–766.

[9] M. S. Yang, C. Y. Lai and C. Y. Lin. "A robust EM clustering algorithm for gaussian mixture models". Patt. Recog., vol. 45, 2012, pp. 3950–3961.

[10] M. S. Yang, and Y. Nataliani. "Robust-learning fuzzy c-means clustering algorithm with unknown number of clusters". Patt. Recog., vol. 71, 2017, pp. 45–59.

[11] S. Chatzis and T. Varvarigou. "Robust fuzzy clustering using mixtures of student's-t distributions". Patt. Recog., vol. 29, 2008, pp. 1901–1905.

[12] Z. Ju and H. Liu. "Fuzzy gaussian mixture models". Patt. Recog. vol. 45, 2012, pp. 1146–1158.

[13] Quost, B., Denoeux, T., 2016. "Clustering and classification of fuzzy data using the fuzzy EM algorithm". Fuz. Set. and Sys., vol. 286, 2016, pp. 134–156.

[14] F. Hashemi, M. Naderi and M. Mashinchi. "Clustering right-skewed data stream via Birnbaum–Saunders mixture models: A flexible approach based on fuzzy clustering algorithm". Appl. Sof. Compu., vol. 82, 2019, 105539.

[15] D. B. Rubin. "Inference and missing data". Biomet. vol. 63, 1976, 581–592.

[16] D. B. Rubin. "Multiple imputation for nonresponse in surveys". vol. 81, 2004, John Wiley & Sons.

[17] M. S. Yang, C. Y. Lai and C. Y. Lin. "A robust EM clustering algorithm for gaussian mixture models". Patt. Recog., vol. 45, 2012, pp. 3950–3961.

[18] T. I. Lin. "Learning from incomplete data via parameterized t mixture models through eigenvalue decomposition". Compu. Stat. & Dat. Anal. vol. 71, 2014, pp. 183–195

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Location of fire station in Bam city using Fuzzy Analytic Hierarchy Process

Amirabbas Abouei Mehrizi[1], Tahereh Kordi Karimabadi[2]

[1]*Department of Industrial Engineering, Higher Education Complex of Bam, Bam, Iran,* abouei@bam.ac.ir
[2]*Master student of Industrial Engineering, Shahid Bahonar University of Kerman, Kerman,* t.kordi7327@gmail.com

***Abstract*—Creating new urban facilities necessitates an in-depth examination of the optimal location of these facilities within a city. To properly allocate urban facilities, the first fundamental step is to determine the optimal location based on a variety and sometimes contradictory conditions. This issue becomes critical when critical factors such as saving human lives are considered. As a result, the optimal location of fire stations is a critical issue, given the value placed on the lives of those at risk. This is especially evident in Bam, which is a garden city. To locate fire stations in this paper, the situation was first examined, and based on the existing stations, the radius of operation for each station was determined using the standard of Bam Standard Institute and population of the city, Number-6430. As a result, the city requires another station. After selecting a few suitable locations, pairwise comparison tables with triangular fuzzy numbers were completed and the proposed locations were prioritized using the expert's opinion. Finally, the optimal location was determined through the use of the fuzzy analytic hierarchy process (FAHP) method.**

***Index Terms* —Location, Fire Station, Bam, Fuzzy Hierarchical Analysis (FAHP)**

I. INTRODUCTION

The prestige and importance of a city are directly related to the services and facilities that support urban settlements. The higher the quality of these services, the more comfortable and affordable it is for citizens to live. If these services were located in suitable and sufficient locations, the economic impact on residents and their time costs would be significantly reduced. This entails conducting scientific and expert research.

One of the main responsibilities of urban planners is to allocate land to various uses in accordance with the city's role and property, the urban economy, and the interaction of land uses. Today, urban planning places a premium on the location of municipal service centers, which include schools, health centers, fire stations, police stations, and chain stores. Urban planners, through their proper placement, also contribute significantly to ensuring citizens' well-being and, as a result, to the realization of social justice. Among the city's existing uses and services, the optimal distribution and location of fire stations is critical, given the importance of safety in cities and the provision of measures to monitor and deal with accidents.

Numerous studies and analyses of fire incidents and fire station operations indicate significant limitations and inadequacies in the location and optimal performance of fire stations. These issues and failures can be classified into the following categories:

- Non-compliance with the location and coverage radius of stations in areas with a high-risk potential.
- Disproportion between the number of stations and the population covered.
- Disproportion between the spatial distribution of stations and the time coverage standard.
- Insufficient number of stations in relation to population and urban area criteria.

The optimal location of service stations has been extensively studied, with numerous models developed in this area. Numerous studies have been conducted to optimize one or more objective functions using a variety of different criteria.

In 2014, Rahnama and Aftab identified fire stations in Urmia based on population densities, appropriate contiguity, inadvertent contiguity, and accessibility. In a geographic information system (GIS), they used the analytic hierarchy process (AHP). According to the findings, 49.2% of the city area is outside the operational area of existing stations, necessitating the construction of four new stations [1].

Khan Ahmadi et al. (2014) located two fire stations in Tehran's region 1 and district 10, using a combination of fuzzy

logic and AHP in the GIS. After establishing the effective factors and criteria, they weighed them using the AHP method and then evaluated the layers using fuzzy logic [2].

Ebrahimi and Mirzaei Madam conducted an investigation in 2015 to determine the location of new emergency services in Tehran. They used the FAHP method to rank and score target areas in their research [3].

Murray conducted a study in 2013 to determine the optimal location of California fire stations. By combining strategic planning, geographic information systems, and the maximum spatial coverage model, he demonstrated that, with the exception of two fire stations, the remaining stations are not in optimal locations [4].

Farhadi and Farhadi conducted a study in 2017 to determine the location of fire stations in Shiraz. They scored and ranked the three target areas using the TOPSIS and AHP methods [5].

In 2020, Nyimbili and Erden developed a model for the optimal location of new fire stations in Istanbul using an integrated approach of FAHP and GIS. To validate the FAHP results, the best-worst method was used, in which the weights of the evaluated criteria were compared and found to be highly correlated; This demonstrates the high level of consistency and reliability of the fuzzy AHP model for the selection of new fire stations and will be beneficial for use as a basis for location planning decisions. Finally, a total of 34 new fire stations were proposed to supplement the existing 121 [6].

Xincong et al. conducted an analysis of the factors and principles governing multi-facility location in 2020. Finally, a new evaluation index system for multi-facility locations was proposed, and the FAHP method was used to thoroughly evaluate candidate locations and choose the best one. The findings indicated that the presented research provides a solid foundation for evaluating the location of several facilities and enhances the scientific and reasonable nature of the location results [7].

Türk and Özkök proposed an integrated fuzzy TOPSIS and FAHP approach to shipyard location in Turkey in 2020. In this study, the relative weights of the primary and sub-criteria were calculated first using FAHP. Then the options were prioritized using fuzzy TOPSIS [8].

A 2020 study by Kaya et al. examined the location of electrical and electronic waste recycling facilities in Istanbul. The Pythagoras FAHP was used to determine the most suitable location among the three potential locations after establishing the relevant criteria and sub-criteria [9].

Guler and Yomralioglu proposed an integrated approach in 2020, utilizing GIS and FAHP techniques, to locate suitable locations for electric vehicle (EV) charging stations in Istanbul. They used FAHP to determine which of the three locations was the best [10].

In 2021, Khademi et al. used a GIS to identify areas in Qazvin County for construction waste disposal, taking into account the terms and conditions of organizations such as the municipality, waste management, and environmental protection organization. The weight of each information layer was then calculated using FAHP, and by combining and overlapping the weighted layers, a map of the location of the Qazvin County construction waste landfill was created [11].

Dhingra et al. examined a variety of factors relevant to locating an electric vehicle charging station in New Delhi's Rohini neighborhood in 2021. They used an integrated AHP and fuzzy TOPSIS method to determine the most suitable location from five possible locations, with the AHP method calculating the weights of the criteria and the fuzzy TOPSIS method ranking the options [12].

In 2021, Kieu et al. proposed a hybrid model for perishable agricultural product distribution location selection based on a spherical fuzzy analytic hierarchy process (SF-AHP) and a combined compromise solution (CoCoSo) algorithm. To demonstrate the efficacy of the proposed method, the distribution of sweet potato crops in Vietnam's Mekong Delta region was examined. The results indicated that the proposed model is applicable to the global agricultural supply chain [13].

II. Fuzzy AHP Model

Multiple-criteria decision making (MCDM) is a process that consists of several steps. Using MCDM techniques to identify issues, prioritize them, evaluate options, and select the best option for decision making is the first step in determining the number of indicators and criteria for the issue.

The next step is to gather the necessary information and data to ensure that the data accurately reflects the decision maker's viewpoints. Then, various options and alternatives will be identified based on them. These options can point the decision-maker in the right direction. Finally, determining the most effective method for evaluating and prioritizing options.

The AHP method is a structured technique that was introduced to the world by Thomas L. Saati in 1972 and has since been widely developed. This method establishes a framework for decomposing complex problems into a logical and simpler hierarchy, which enables the planner to easily evaluate the alternatives in terms of the criteria and sub-criteria. Professor Lotfi A. Zadeh proposed the "fuzzy theory" for quantifying phenomena in 1965. His initial objective was to develop a more efficient model for describing the natural language processing procedure [14].

Saati extended the process of hierarchy analytic in fuzzy space in 1992, following the introduction of fuzzy logic to the sciences of mathematics and engineering. The concept of fuzziness is used to determine the pairwise comparison matrices (PCM) in this method, and fuzzy numbers are used to express the degree of preference of the elements. Following Saati's introduction of the FAHP method, and in light of the critical nature of using fuzzy concepts in decision making, this method has been developed by a number of researchers in recent years; in this paper, the method presented by Chang is used. This method is more straightforward than other FAHP techniques and is similar to the classical AHP technique. Triangular fuzzy numbers are used in this method to avoid ambiguity caused by uncertainty in decision making at all stages. A triangular fuzzy number indicated by the symbol m = (l, m, u). The following are the steps in the FAHP by Chang development analysis method [14]:

Step 1: The creation of a hierarchical structure for the problem

To construct hierarchical levels, it is necessary to identify distinct and related levels between the components of each level and their higher and lower levels. In the simplest case, the first level is concerned with the objective, the second with the desired criteria, and the third with the options under consideration. The levels of a simple hierarchy are depicted in Fig. 1.

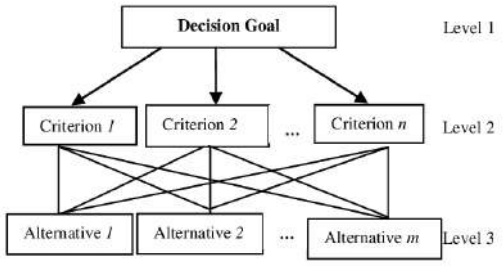


Fig. 1. Simple hierarchy levels.

Step 2: Utilization of fuzzy numbers to generate pairwise comparison matrices

After establishing the hierarchical structure, the pairwise comparison matrices should be determined based on the decision maker's opinion. This operation is repeated for each level that is to be executed. By comparing options, *i* and *j*, the value of *i* over *j* is determined using one of the fuzzy numbers in the Table I:

Step 3: Calculation of the fuzzy synthetic extent

Following the formation of a pairwise comparison matrix with fuzzy numbers, the value of the fuzzy compound expansion (S_i) for each row is calculated. S_i are triangular fuzzy numbers derived from Equation (1):

$$S_i = \sum_{j=1}^{m} \widetilde{M}_{gi}^{j} \otimes \left(\sum_{i=1}^{n} \sum_{j=1}^{m} \widetilde{M}_{gi}^{j} \right)^{-1} \quad (1)$$

Where, $\widetilde{M}_{gi}^{j} = (l_{ij}, m_{ij}, u_{ij})$ are fuzzy numbers of pairwise comparison matrix.

Step 4: Calculation of the degree of preference of S_i (over each other)

The degree of preference for S_i over S_j is calculated in this step using equation (2):

$$V(\tilde{S}_i > \tilde{S}_k) = \begin{cases} 1 & m_i > m_k \\ 0 & l_k > u_i \\ \dfrac{l_k - u_i}{(m_i - u_i) - (m_k - l_k)} & otherwise \end{cases} \quad (2)$$

Step 5: Calculation of the relative weight vector of criteria and options

Equation (3) is used to determine the relative weights of the criteria in relation to the objective and the options in relation to each of the criteria:

$$\acute{d}(A_i) = \min V(S_i \geq S_k) \quad \forall\, k = 1,2,\dots,n \quad (3)$$

In this case, $\acute{d}(A_i)$ denotes the relative weight. As expression (4): for each matrix, a relative weight vector is obtained:

$$\acute{W} = (\acute{d}(A_1), \acute{d}(A_2), \dots, \acute{d}(A_n)) \quad (4)$$

TABLE I. HOW TO CONVERT LINGUISTIC EXPLANATION TO FUZZY NUMERICAL LOGIC.

linguistic explanation	fuzzy numerical logic	linguistic explanation	fuzzy numerical logic
Quite a lot of preference	(8,9,10)	Medium to high preference	(3,4,5)
Preference is very high to quite high	(7,8,9)	Medium preference	(2,3,4)
Too much preference	(6,7,8)	Low to medium preference	(1,2,3)
High preference to very high	(5,6,7)	Same preference	(1,1,1)
High preference	(4,5,6)		

Where, $\acute{W}$ is a vector of relative weights. The normalized vector is then obtained by dividing each entry of the relative weight vector by the sum of the vector's total weights.

Step 6: Calculation of the final weight

The final weight assigned to each option is calculated by multiplying the rate of importance of the criteria by the weight of the options.

Calculation of the inconsistency ratio

The consistency or inconsistency of the pairwise comparison matrix of the criteria is the material that entirely affects the course of the study in the research of hierarchical structures. Its critical to determine the inconsistency ratio (I.R) during the initial validation of the paired comparison data and apply it in decision-making, since if the inconsistency ratio is more than 0.1, the pairwise comparison matrix will have very little validity, putting the decision in doubt. The inconsistency index (I.I) must be obtained before the inconsistency ratio can be calculated. Using Equation (5) and the center of area (CoA), the fuzzy numbers are converted to absolute numbers to compute the inconsistency ratio in fuzzy matrices [14]. Then, using equations (6) to (9), we complete the appropriate calculations.

$$a_{ij} = \left(\frac{l_{ij}, m_{ij}, u_{ij}}{3}\right) \quad (5)$$

$$\lambda_{\text{maxi}} = \frac{\text{A} \times \text{W}}{\text{W}_i} \quad (6)$$

$$\lambda_{max} = \frac{\sum_{k=0}^{n} \lambda_{maxi}}{\mathrm{n}} \quad (7)$$

$$\mathrm{I.I} = \frac{\lambda_{\max} - \mathrm{n}}{\mathrm{n} - 1} \quad (8)$$

$$\mathrm{I.R} = \frac{\mathrm{I.I}}{\mathrm{I.I.R}} \quad (9)$$

Where, a_{ij} is the absolute value of the components of the pairwise comparison matrix, A is the pairwise comparison matrix, W is the weights matrix, I.I is the inconsistency index, I.R is the inconsistency ratio, and I.I.R is the inconsistency index of random.

The value of the inconsistency index of random matrix (I.I.R). is examined in Table II to determine the inconsistency ratio, according to the dimension of the matrix:

III. STATION QUO AND CASE STUDY REGION

Bam is the capital of Kerman province. Kerman province is also in Iran's southeast. According to the 2016 Iranian Statistics Center census, Bam had a population of 318241 people. This city is bounded on the north and west by Kerman, on the southwest by Jiroft, on the south by Anbarabad, and on the east by Narmashir and Mohammadabad. Bam is located 200 km southeast of Kerman on the Kerman-Zahedan road.

TABLE II. THE VALUE OF THE INCONSISTENCY INDEX ACCORDING TO THE DIMENSIONS OF THE MATRIX.

N	1	2	3	4	5	6	7	8
I.I.R	0	0	0.58	0.9	1.12	1.24	1.32	1.41

The general policy of establishing fire stations in Iran, has been unplanned and uncoordinated. To establish any station in urban areas, the most critical principle is to avoid using land without first owning it or relying on other factors that diminish the value of land, which has influenced the location of fire stations in cities. This was also true for the location of the Bam fire stations. This city has two fire stations, one at Attarzadeh junc and another on ShahidRajaei Blvd., as illustrated in Fig. 2.

According to international standards, the time allowed between the start of a fire and the start of a firefighting operation is 3 to 5 minutes. Because, according to the standard time-temperature curve, if firefighters are unable to arrive on the scene within the first five minutes of a fire and attempt to extinguish it, the fire will spread and have dire consequences. Additionally, small stations' operating radius is limited to 2 km by fire station location rules. Naturally, in high-fire-risk areas and locations with unique problems, such as the chemical industry, warehouses, markets, commercial complexes, economic centers, and cities, a working radius of 1.6 km should be considered [15].

As illustrated in Fig. 2, the western parts of the city are beyond the operational limits of existing stations, and based on the operational radius of existing stations, the population of Bam, and the existing stations, the city requires another fire station. In order to locate this station, we first suggest a few places to build a fire station. The locations proposed are depicted in Fig. 3. Then, the best location is determined using pairwise comparison tables based on expert opinions and the FAHP method.

A. Determining the criteria

To locate fire stations, one must first determine the effective criteria for locating them. The criteria in this paper are defined in accordance with the criteria specified in standard Number-6430, and other projects, papers, and researches conducted in this field, as well as expert opinion. The criteria and their significance are detailed below.

Fig. 2. Fire station in Bam.

Fig. 3. Locations suggested for the construction of a fire station.

1) Density

The population density of a region is a significant factor in determining the location of the stations and associated planning. The probability of fire is greater in areas of the city with a high population density than in areas with a low population density. As a result, fire centers should be strategically located in accordance with the population density of various urban areas.

2) Accessibility

It is critical to avoid selecting unsuitable locations when locating stations (inappropriate slopes, inappropriate elevation, and abnormal surface water conditions). When determining a location, care must be taken to ensure that the site chosen has the fewest available and future access barriers that increase the time required to reach the scene. Along with the availability of arterial route networks, it is necessary to provide access to water, electricity, and telephone networks, and transportation systems when locating and establishing fire stations [15].

3) Risk potential

Risk assessment in various urban areas is based on the number and frequency of accidents, identifying vulnerable points in fire accidents and locations with high-risk potential. Station locations should take precedence over such areas.

4) Proximity to the fuel station

Proximity to a variety of urban uses improves the quality and efficiency of urban service centers, including fire stations. Due to the presence of combustible materials and a higher fire risk in this study, the fuel station is considered a compatible use. It is possible for the stations to be located closest to this user.

5) Being away from religious institutions, schools, and hospitals

Due to the impact of fire stations on people's quality of life, it is preferable to avoid placing them near incompatible uses such as hospitals, religious centers, and schools, as these locations experience peak traffic during the day. As a result, distance restrictions must be imposed on the station's proximity to such uses. According to standard Number-6430, the fire station's minimum distance from these locations is 500 m, and the greater the distance, the better.

6) Operating radius

It is essential to evaluate the qualitative and quantitative relief performance of each station at the time of the accident. Thus, the radius of operation of each station is considered in order to expeditiously reach the accident site, provide relief, increase station coverage, and mitigate financial and human risks associated with the accident.

B. Hierarchical structure of problem

The hierarchical structure of the problem is shown in Fig. 4.

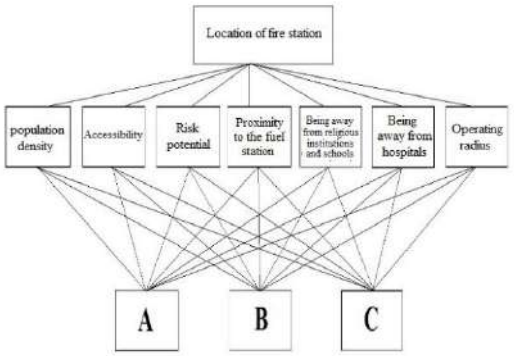


Fig. 4. Hierarchical structure of locating the fire station.

C. Pairwise comparison matrices and check data consistency

The pairwise comparison matrix was completed based on the expert's opinion, and the data inconsistency ratio was calculated using the center of area method in Expert Choice software after absolute fuzzy numbers. The inconsistency ratio of the data is acceptable. The findings are summarized in Tables III to X. Each table's lower left column indicates the inconsistency ratio.

TABLE III. PAIRWISE COMPARISON MATRIX OF POPULATION DENSITY CRITERION.

C1	**A**	**B**	**C**
A	(1,1,1)	(5,6,7)	(1,2,3)
B	(0.14,0.17,0.2)	(1,1,1)	(0.14,0.17,0.2)
C	(0.33,0.5,1)	(5,6,7)	(1,1,1)
0.05			

TABLE IV. PAIRWISE COMPARISON MATRIX OF ACCESSIBILITY CRITERION.

C2	**A**	**B**	**C**
A	(1,1,1)	(3,4,5)	(1,2,3)
B	(0.2,0.25,0.33)	(1,1,1)	(0.25,0.33,0.5)
C	(0.33,0.5,1)	(2,3,4)	(1,1,1)
0.02			

TABLE V. PAIRWISE COMPARISON MATRIX OF RISK POTENTIAL CRITERION.

C3	**A**	**B**	**C**
A	(1,1,1)	(0.33,0.5,1)	(3,4,5)
B	(1,2,3)	(1,1,1)	(4,5,6)
C	(0.2,0.25,0.33)	(0.17,0.2,0.25)	(1,1,1)
0.01			

TABLE VI. PAIRWISE COMPARISON MATRIX OF PROXIMITY TO THE FUEL STATION CRITERION.

C4	**A**	**B**	**C**
A	(1,1,1)	(3,4,5)	(1,2,3)
B	(0.2,0.25,0.33)	(1,1,1)	(0.25,0.33,0.5)
C	(0.33,0.5,1)	(2,3,4)	(1,1,1)
0.02			

TABLE VII. PAIRWISE COMPARISON MATRIX OF BEING AWAY FROM RELIGIOUS INSTITUTIONS AND SCHOOLS, CRITERION.

C5	**A**	**B**	**C**
A	(1,1,1)	(0.33,0.5,1)	(3,4,5)
B	(1,2,3)	(1,1,1)	(4,5,6)
C	(0.2,0.25,0.33)	(0.17,0.2,0.25)	(1,1,1)
0.02			

TABLE VIII. PAIRWISE COMPARISON MATRIX OF BEING AWAY FROM HOSPITALS CRITERION.

C6	**A**	**B**	**C**
A	(1,1,1)	(1,1,1)	(1,1,1)
B	(1,1,1)	(1,1,1)	(1,1,1)
C	(1,1,1)	(1,1,1)	(1,1,1)
0			

TABLE IX. PAIRWISE COMPARISON MATRIX OF OPERATING RADIUS CRITERION.

C7	A	B	C
A	(1,1,1)	(1,2,3)	(3,4,5)
B	(0.33,0.5,1)	(1,1,1)	(2,3,4)
C	(0.2,0.25,0.33)	(0.25,0.33,0.5)	(1,1,1)
0.02			

TABLE X. PAIRWISE COMPARISON MATRIX OF CRITERION.

	C1	C2	C3	C4	C5	C6	C7
C1	(1,1,1)	(1,2,3)	(0.33,0.5,1)	(1,1,1)	(3,4,5)	(1,2,3)	(0.33,0.5,1)
C2	(0.33,0.5,1)	(1,1,1)	(1,1,1)	(2,3,4)	(4,5,6)	(1,2,3)	(1,1,1)
C3	(1,2,3)	(1,1,1)	(1,1,1)	(2,3,4)	(5,6,7)	(1,1,1)	(1,1,1)
C4	(1,1,1)	(0.25,0.33,0.5)	(0.25,0.33,0.5)	(1,1,1)	(3,4,5)	(0.33,0.5,1)	(0.33,0.5,1)
C5	(0.2,0.25,0.33)	(0.17,0.2,0.25)	(0.14,0.17,0.2)	(0.2,0.25,0.33)	(1,1,1)	(0.14,0.17,0.2)	(0.11,0.13,0.14)
C6	(0.33,0.5,1)	(0.33,0.5,1)	(1,1,1)	(1,2,3)	(5,6,7)	(1,1,1)	(1,1,1)
C7	(1,2,3)	(1,1,1)	(1,1,1)	(1,2,3)	(7,8,9)	(1,1,1)	(1,1,1)
0.05							

D. Calculation of the relative weight of options and criteria

The relative weight of the criteria in relation to the options and the options in relation to the objective were calculated using Chang's development analysis method; results are shown in Tables XI and XII.

E. Calculation of the final weight of the options

Additionally, the final weight assigned to each option indicates its rank. The greater an option's final weight, the higher and better the rank. The final weights assigned to the options are listed in Table XIII. According to the results, the highest-weighted location **A** was given precedence. Following that, options **C** and **B** are ranked second and third.

IV. CONCLUSION

Urban relief centers are one of the physical components required by urban settlements, and their importance is growing daily from a variety of perspectives as cities expand. The location of relief stations is critical in ensuring the city's security and development. Fire stations, as waiting areas for fire trucks and ambulances, are critical service centers in cities, contributing significantly to the safety and comfort of citizens and the economic development of cities. Furthermore, a city like Bam, which is a garden city, is more vulnerable to threats. The critical nature of these uses necessitates the use of the proper process for locating these stations.

The results of the research and review of existing stations, as well as the operating radius of each station in accordance with the standard Number-6430 of the Institute of Standards and Population of Bam, indicate that the western part of the city is not served by existing stations, which explains why there are not enough stations to cover the population. It encompasses the entire city and requires the location and construction of one new station. As a result, the need for careful site selection and the construction of a new fire station to cover the area is justified.

To determine the most suitable location for the fire station, three potential locations were initially introduced and examined. Obviously, it is critical to identify the most effective criteria for this selection and the extent to which each of the criteria impacts the outcome. The FAHP technique was used to rank and score the options, as it is a versatile and robust decision-making technique.

TABLE XI. RELATIVE WEIGHT OF THE CRITERIA IN RELATION TO THE OBJECTIVE.

criteria	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**
Relative weight	0.148	0.191	0.216	0.051	0	0.160	0.234

TABLE XII. RELATIVE WEIGHT OF THE OPTIONS IN RELATION TO EACH CRITERION.

	C1	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**
A	0.564	0.616	1	0.184	0.384	0.333	0.616
B	0	0	0	0	0.616	0.333	0.384
C	0.436	0.384	0	0.816	0	0.333	0

TABLE XIII. FINAL WEIGHT OF THE OPTIONS.

Options	**weight**	**Rank**
A	0.626	1
B	0.142	3
C	0.232	2

According to the findings of the study on locating the fire station, it is observed that the station's location is prioritized according to the following criteria: In the first place, **A** is located on Sahebolzaman Blvd. In the second place, **C** is located at the beginning of ShahidBahonar Blvd., and in the third place, **B** is located on the Khalij Fars highway.

REFERENCES

[1] M. R. Rahnama, and A. Aftab, "Location of fire stations in Urmia using GIS and AHP," Journal of Geography and Development. No. 35. pp.166-153. 1393, (in persian).

[2] M. Khan Ahmadi, M. Arabi, A. Vafainejad and H. Rezaian, "Location of fire stations using a combination of Fuzzy and AHP logic in GIS environment (Case study: District 1, District 10 of Tehran)," Quarterly Journal of Geographical Information Research "Sepehr", Volume 23, Issue 89, pp. 88-98. 1393, (in persian).

[3] M. Ebrahimi and M. Mirzaei Madam, "Ranking of Tehran regions for adding new emergency services by fuzzy AHP method," Journal of Industrial Engineering, Volume 49, Issue 2, pp. 149-163, Faculty of Engineering, University of Tehran, 2015, (in persian).

[4] A. T. Murray, "Optimizing the Spatial Location of Urban Fire Stations", Fire safety journal. 2013.

[5] Z. Farhadi and M. Farhadi, "Evaluation of location of fire stations in Shiraz by AHP and TOPSIS techniques," International Conference on Civil Engineering, Architecture and Urbanism of Contemporary Iran, Tehran, 2017, (in persian).

[6] P. H. Nyimbili and T. Erden, "GIS-based fuzzy multi-criteria approach for optimal site selection of fire stations in Istanbul,

Turkey," Socio-Economic Planning Sciences, Vol. 71, No. 10086, 2020.

[7] L. Xincong, Z. Mingze and Q. Cheng, "Location Evaluation of Multi-station Integration Based on AHP-Fuzzy Comprehensive Evaluation Method," 12th IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC) , p. 1-5, 2020.

[8] A. Türk and M. Özkök, "Shipyard location selection based on fuzzy AHP and TOPSIS," Journal of Intelligent & Fuzzy Systems, Vol. 39, No. 3, 2020.

[9] A. Kaya, B. Çiçekalan and F. Çebi, "Location selection for WEEE recycling plant by using Pythagorean fuzzy AHP," Journal of Intelligent & Fuzzy Systems, Vol. 38, No. 1, 2020.

[10] D. Guler and T. Yomralioglu, "Suitable location selection for the electric vehicle fast charging station with AHP and fuzzy AHP methods using GIS", Annals of GIS, Vol. 26, 2020.

[11] M. Khademi Shiraz, M. Ravanshadnia, A. Khoshand and H. Abbasianjahromi, "Site Selection of Construction Waste Landfill Based on Combination of Fuzzy AHP and Geospatial Information System (GIS)(Case Study: Qazvin, Iran)," Amirkabir Journal of Civil Engineering, Vol. 53, No.7, 2021.

[12] Dhingra, A., Jareda, A., Choudhary, H., Agrawal, S.*" Selection of optimal electric vehicle charging station location using AHP-fuzzy TOPSIS approach",* Proceedings of the 2nd International Conference on ICT for Digital, Smart, and Sustainable Development, ICIDSSD, 2020.

[13] P. T. Kieu, V. T. Nguyen and T. P. Ho, "A Spherical Fuzzy Analytic Hierarchy Process (SF-AHP) and Combined Compromise Solution (CoCoSo) Algorithm in Distribution Center Location Selection: A Case Study in Agricultural Supply Chain," Axioms Journal, Vol. 10, No. 53, 2021.

[14] M. Amiri, A. Darestani Farahani and M. Mahboub Ghodsi, Multi-Criteria Decision Making, edited by Mohammad Aliha, Hamid, first edition, Tehran, Kian University Press, 2016, (in persian).

[15] Iranian Institute of Standards and Industrial Research, Standard 6430 (Rules for the location criteria of urban fire stations), first edition, 2001, (in persian).

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Minimizing a composite fuzzy function in terms of subgradient

1st Samira Fatemi, 2nd Ildar Sadeqi and 3rd Fridoun Moradlou

1st Department of Mathematics, *Sahand University of Technology*, Tabriz, Iran, s_fatemi98@sut.ac.ir
2nd Department of Mathematics, *Sahand University of Technology*, Tabriz, Iran, esadeqi@sut.ac.ir
3rd Department of Mathematics, *Sahand University of Technology*, Tabriz, Iran, moradlou@sut.ac.ir

***Abstract*— The purpose of this paper is expression of the concepts of directional derivative, gradient and subdifferential of fuzzy mappings from $\mathbb{R}^n$ into $\mathcal{F}_0$. Moreover, we study optimality conditions for the minimizing problem for the composite of fuzzy mappings.**

***Index Terms*— Fuzzy number, Fuzzy mapping, Directional derivative, Differential, Subdifferential, Minimizing a composite fuzzy mapping.**

I. Introduction

In many scientific areas, such as systems analysis and operations research, a model has to be set up using data which is only approximately known. Fuzzy sets theory, introduced by Zadeh [1], makes this possible. In 1972, Chang and Zadeh [2] introduced the concept of fuzzy numbers with the consideration of the properties of probability functions. The fuzzy numbers have been extensively studied by many authors. With the development of theories of fuzzy numbers and its applications the concept of fuzzy numbers becomes more and more important.

The theory of fuzzy differential calculus has been discussed by many researchers like Goetschel–Voxman [3], Seikkala [4], Puri–Ralescu [5], Dubois–Prade [6, 7], and Friedman–Ming–Kandel [8]. A comparison of these various definitions has been discussed by Buckley–Feuring[9]. Goetschel–Voxman, Puri–Ralescu, and Friedman–Ming–Kandel have used non-standard fuzzy subtraction to define derivative of a fuzzy mapping. Buckley–Feuring [9, 10] also have defined the derivative of a fuzzy mapping using left and right-hand functions of its α-level sets and established sufficient conditions for the existence of fuzzy derivative. Panigrahi, Panda and Nanda [15] have defined the gradient of a fuzzy function using left and right-hand functions of its α-level sets. Wang and Wu [14] have defined the concepts of some kinds of directional derivative of fuzzy mappings from $\mathbb{R}^n$ into $\mathcal{F}_0$, and discussed the problems of characterizations of the directional derivatives and existence of directional derivative for convex fuzzy mappings. Also they have defined the concepts of differential which is different from the de Blasi differential (Diamond and Kloeden defined in [16]) and subdifferential, discussed the relation between gradients and partial derivatives, and presented the characterization of differentiability for fuzzy mapping.

Nanda and Kar [11] introduced the concept of convexity for fuzzy mapping and discussed the concepts of convex fuzzy mappings from a vector space over the field $\mathbb{R}^n$ into $\mathcal{F}_0$, established criteria for convex fuzzy mappings. In 1999, Yu-Ru Syau [12] also established two characterizations for convex fuzzy mapping.

Since the set of fuzzy numbers is a partially ordered set, it might happen that two fuzzy numbers may not be comparable. Thus, in such case one is not sure what is the maximum or minimum of two fuzzy numbers (that is, when they are not comparable). So to overcome this difficulty, Syau [12] has taken a different approach and defined the supremum and infimum for a pair of fuzzy numbers. In this paper, we consider this definition of infimum for the problem of minimization of a fuzzy mapping.

If the values of the target function that is sought optimum solution are crisp real numbers, the programming is a general crisp programming. But in reality, sometimes, the values of the target function only are estimated values, so it is more suitable that the values are expressed with fuzzy numbers. In 1997, Butnariu [13] studied the problem of the fuzzy convex programming for convex continuous fuzzy mapping from $\mathcal{F}_0(\mathbb{R}^n)$ into $\mathbb{R}$ and presented an iterative method of solving the convex optimization problems.

In section (III), we will present a necessary optimality condition for problems consisting of minimizing a composite fuzzy function $f + g$, where f possesses some differentiability

properties but is not assumed to be convex while g is a convex fuzzy mapping but is not assumed it have any special differentiability properties.

II. Preliminaries

We now quote some definitions which will be needed in the sequel.

Definition II.1. *Let $\mathbb{R}$ denote the set of all real numbers. A fuzzy number is a mapping $u : \mathbb{R} \to [0,1]$ with the following properties:*

(1) *u is normal, that is, there exists an $x_0 \in \mathbb{R}$ with $u(x_0) = 1$;*
(2) *u is a fuzzy convex fuzzy set, that is, $u((1-\alpha)x+\alpha y) \geqslant \min(u(x), u(y))$ whenever $x; y \in \mathbb{R}$ and $\alpha \in [0,1]$;*
(3) *$u(x)$ is upper semi-continuous;*
(4) *$[u]^0 = \{x \in \mathbb{R} : u(x) > 0\}$ is a compact set.*

Let $\mathcal{F}_0$ be the set of all fuzzy numbers on $\mathbb{R}$. The α-level set of a fuzzy number $u \in \mathcal{F}_0$, $0 \leq \alpha \leq 1$, denoted by $[u]^\alpha$, is defined as

$$[u]^\alpha = \begin{cases} [u]^0 = \{x \in \mathbb{R} : u(x) \geq \alpha\} & \text{if } 0 < \alpha \leq 1, \\ \text{cl}(\text{suppu}) & \text{if } \alpha = 0. \end{cases}$$

It is clear that the α-level set of a fuzzy number is a closed and bounded interval $[u_*(\alpha), u^*(\alpha)]$ where $u_*(\alpha)$ denotes the left-hand end point of $[u]^\alpha$ and $u^*(\alpha)$ denotes the right-hand endpoint of $[u]^\alpha$. Also any $m \in \mathbb{R}$ can be regarded as a fuzzy number $\hat{m}$ defined by

$$\hat{m}(t) = \begin{cases} 1 & \text{if } t = m, \\ 0 & \text{if } t \neq m. \end{cases}$$

In particular, the fuzzy number $\hat{0}$ is defined as $\hat{0}(t) = 1$ if $t = 0$, and $\hat{0}(t) = 0$ otherwise. Thus a fuzzy number u can be identified by a parameterized triples

$$\{(u_*(\alpha), u^*(\alpha), \alpha) : \alpha \in [0,1]\}.$$

A linear structure in $\mathcal{F}_0$ is defined via the following operations:

$$(u+v)(x) = \sup_{y+z=x} \min[u(y), v(z)],$$

$$(\lambda u)(x) = \begin{cases} u(\lambda^{-1}x) & \text{if } \lambda \neq 0, \\ \hat{0} & \text{if } \lambda = 0 \end{cases}$$

for $u, v \in \mathcal{F}_0$, $\lambda \in \mathbb{R}$.

It is well known that for any $u, v \in \mathcal{F}_0$ and $\lambda \in \mathbb{R}$, $u+v$, $\lambda u \in \mathcal{F}_0$ and $[u+v]^\alpha = [u]^\alpha + [v]^\alpha$, $[\lambda u]^\alpha = \lambda[u]^\alpha$, i.e.

$$(u+v)_*(\alpha) = u_*(\alpha)+v_*(\alpha), \quad (u+v)^*(\alpha) = u^*(\alpha)+v^*(\alpha)$$

and

$$(\lambda u)_*(\alpha) = \begin{cases} \lambda u_*(\alpha) & \text{if } \lambda \geq 0 \\ \lambda u^*(\alpha) & \text{if } \lambda < 0 \end{cases}$$

$$(\lambda u)^*(\alpha) = \begin{cases} \lambda u^*(\alpha) & \text{if } \lambda \geq 0 \\ \lambda u_*(\alpha) & \text{if } \lambda < 0 \end{cases}$$

for every $\alpha \in [0,1]$.

for $x = (x_1, x_2, ..., x_n), y = (y_1, y_2, ..., y_n) \in \mathbb{R}^n$, we define $x \leq y$ iff $x_i \leq y_i (i = 1, 2, ..., n)$. And we define $x < y$ iff $x \leq y$ and $x \neq y$.

for $u, v \in \mathcal{F}_0$, we define a partial ordering $\leq$ in $\mathcal{F}_0$ by $u \leq v$ iff $[u]^\alpha = [u_*(\alpha), u^*(\alpha)] \leq [v]^\alpha = [v_*(\alpha), v^*(\alpha)]$ for any $\alpha \in [0,1]$, and $[u]^\alpha \leq [v]^\alpha$ iff $u_*(\alpha) \leq v_*(\alpha)$ and $u^*(\alpha) \leq v^*(\alpha)$. Also, we define $u < v$ iff $u \leq v$ and $u \neq v$.

Lemma II.2. *[18] If $u, v, w, h \in \mathcal{F}_0$, then*

(1) $u \leq v \Longleftrightarrow u + w \leq v + w$;
(2) $u \leq v \Rightarrow ku \leq kv, \mu u \geq \mu v, (k \geq 0, \mu < 0)$;
(3) $u \leq v \Longleftrightarrow -u \geq -v$;
(4) $u \leq v, w \leq h \Rightarrow u + w \leq v + h$,
$u < v, w \leq h \Rightarrow u + w < v + h$;
(5) $v = w \Longleftrightarrow u + v = u + w$;
(6) $u \leq v, w \leq h, u + w = v + h \Rightarrow u = v, w = h$.

Theorem II.3. *(Representation theorem [14, 15])*
If $u \in \mathcal{F}_0$, then $u_* : [0,1] \to \mathbb{R}$ and $u^* : [0,1] \to \mathbb{R}$ satisfy the conditions:

(1) *$u_* : [0,1] \to \mathbb{R}$ is a bounded increasing function,*
(2) *$u^* : [0,1] \to \mathbb{R}$ is a bounded decreasing function,*
(3) *$u_*(1) \leq u^*(1)$ (it is equivalent to $u_*(\alpha) \leq u^*(\alpha)$ for any $\alpha \in [0,1]$),*
(4) *for $0 < k \leq 1$, $\lim_{\alpha \to k^-} u_*(\alpha) = u_*(k)$ and $\lim_{\alpha \to k^-} u^*(\alpha) = u^*(k)$,*
(5) *$\lim_{\alpha \to 0^+} u_*(\alpha) = u_*(0)$ and $\lim_{\alpha \to 0^+} u^*(\alpha) = u^*(0)$.*

Conversely, if the functions $u_*(\alpha)$ and $u^*(\alpha)$ on $[0,1]$ satisfy conditions (1)–(5), then there exists a unique $u \in \mathcal{F}_0$ such that $[u]^\alpha = [u_*(\alpha), u^*(\alpha)]$ for each $\alpha \in [0,1]$.

$\tilde{A}$ is said to be an n-dimensional fuzzy vector if the components of $\tilde{A}$ are composed by n fuzzy numbers, which is denoted by $\tilde{A} = (x_1, x_2, ..., x_n)^T$. The set of all n-dimensional fuzzy vectors is denoted by $\mathcal{F}_0(\mathbb{R}^n)$.

A α-level vector of fuzzy vector $\tilde{A} = (x_1, x_2, ..., x_n)^T$ is defined as

$$[\tilde{A}]_\alpha := ([x_1]_\alpha, [x_2]_\alpha, ..., [x_n]_\alpha)^T$$

and

$$\tilde{A}_*(\alpha) := (x_{1_*}(\alpha), x_{2_*}(\alpha), ..., x_{n_*}(\alpha))^T,$$
$$\tilde{A}^*(\alpha) := (x_1^*(\alpha), x_2^*(\alpha), ..., x_n^*(\alpha))^T,$$

for any $\alpha \in [0,1]$.

The addition and the scalar multiplication of fuzzy vectors $\tilde{A} = (x_1, x_2, ..., x_n)^T$ and $\tilde{B} = (x_1, x_2, ..., x_n)^T$ are defined as

$$\tilde{A} + \tilde{B} := (x_1 + y_1, x_2 + y_2, ..., x_n + y_n)^T,$$
$$k\tilde{A} := (kx_1, kx_2, ..., kx_n)^T (k \in \mathbb{R}).$$

Obviously, if $k > 0$, then

$$kx_i = k\{(x_{i_*}(\alpha), x_i^*(\alpha), \alpha) : \alpha \in [0,1]\}$$
$$= \{(kx_{i_*}(\alpha), kx_i^*(\alpha), \alpha) : \alpha \in [0,1]\},$$

$i = 1, 2, ..., n$.

Let $\tilde{A} = (x_1, x_2, ..., x_n)^T, \tilde{B} = (x_1, x_2, ..., x_n)^T \in \mathcal{F}_0(\mathbb{R}^n)$, then the inner product of fuzzy vectors $\tilde{A}$ and $\tilde{B}$, denoted by $\langle \tilde{A}, \tilde{B} \rangle$, is defined as

$$x_1.y_1 + ... + x_n.y_n,$$

which is a fuzzy number.

Let $\mathbb{R}^n_+ = \{x = (x_1, x_2, ..., x_n)^T : x_i \geq 0 (i = 1, 2, ..., n)\}$, $\tilde{A} = (x_1, x_2, ..., x_n)^T \in \mathcal{F}_0(\mathbb{R}^n)$, and $c \in \mathbb{R}^n_+$, then $c.\tilde{A} \in \mathcal{F}_0$.

Definition II.4. *[19] The Hausdorff distance between two fuzzy vectors is the function $D : \mathcal{F}_0(\mathbb{R}^n) \times \mathcal{F}_0(\mathbb{R}^n) \to [0, \infty)$, defined in terms of the Hausdorff distance between their level sets, that is,*

$$D(u, v) = \sup_{\alpha \in [0,1]} d_H([u]^\alpha, [v]^\alpha) \quad (1)$$

$$= \sup_{\alpha \in [0,1]} \max(|u_*(\alpha) - v_*(\alpha)|, |u^*(\alpha) - v^*(\alpha)|), \quad (2)$$

for any $u, v \in \mathcal{F}_0(\mathbb{R}^n)$.

$(\mathcal{F}_0, D)$ is a complete metric space, and satisfies

$$D(u, v) = D(u + w, v + w), D(ku, kv) = |k| D(u, v)$$

for any $u, v, w \in \mathcal{F}_0$, $k \in \mathbb{R}$.

Let S be a nonempty set in $\mathbb{R}^n$, a mapping $f : S \to \mathcal{F}_0$ is said to be a fuzzy mapping. For any $\alpha \in [0, 1]$, denote $[f(x)]^\alpha = [f(x)_*(\alpha), f(x)^*(\alpha)]$.

A subset C of $\mathbb{R}^n$ is said to be convex if $((1-\lambda)x + \lambda y) \in C$ whenever $x, y \in C$ and $\lambda \in (0, 1)$.

Theorem II.5. *[11, 12] fuzzy mapping $f : C \to \mathcal{F}_0$ defined on a convex subset C in $\mathbb{R}^n$ is convex (respectively, concave) if and only if*

$$f((1 - \lambda)x + \lambda y) \leq (\textit{respectively}, \geq)(1 - \lambda) f(x) + \lambda f(y),$$

for every $x, y \in C$ and $\lambda \in [0, 1]$.

According to the parametric representation of fuzzy number, a fuzzy mapping $f(x)$ can be written as follows:

$$f(x) = \{(f(x)_*(\alpha), f(x)^*(\alpha), \alpha) : \alpha \in [0, 1])\}.$$

Theorem II.6. *[18] Let C be a convex subset in $\mathbb{R}^n$ and $f : C \to \mathcal{F}_0$ be a fuzzy mapping, then f is convex if and only if $f(x)_*(\alpha)$ and $f(x)^*(\alpha)$ are a11 convex functions of x for any fixed $\alpha \in [0, 1]$.*

Theorem II.7. *[18] Let C be a convex subset in $\mathbb{R}^n$, $f : S \subseteq \mathbb{R}^n \to \mathcal{F}_0$ be a convex fuzzy mapping, then fuzzy vector $u \in \mathcal{F}_0(\mathbb{R}^n)$ is a subgradient of f at $x \in C$ if and only if $u_*(\alpha)$ and $u^*(\alpha)$ are subgradients of convex functions $f(x)_*(\alpha)$ and $f(x)^*(\alpha)$ at $x \in C$ for any fixed $\alpha \in [0, 1]$, respectively.*

III. Directional derivative and subdifferential

Definition III.1. *(Directional derivative of a fuzzy function[14]). Let $f : S \subseteq \mathbb{R}^n \to \mathcal{F}_0$ be a fuzzy mapping. For $x \in S$, let $d \in \mathbb{R}^n$ such that $x + \lambda d \in S$ for $\lambda > 0$ and sufficiently small. The directional derivative of f at x along the vector d (if it exists) is a fuzzy number denoted by $f'(x; d)$ and whose α-level set is defined as,*

$$[f'(x; d)]^\alpha = [f'(x; d)_*(\alpha), f'(x; d)^*(\alpha)],$$

where $f'(x; d)_*(\alpha) = \lim_{\lambda \to 0^+} \frac{f(x+\lambda d)_*(\alpha) - f(x;d)_*(\alpha)}{\lambda}$

and $f'(x; d)^*(\alpha) = \lim_{\lambda \to 0^+} \frac{f(x+\lambda d)^*(\alpha) - f(x;d)^*(\alpha)}{\lambda}$.

Definition III.2. *An n-dimensional fuzzy vector u is said to be a subgradient of a convex fuzzy mapping $f : S \subseteq \mathbb{R}^n \to \mathcal{F}_0$ at $x \in C$ if*

$$f(z) \geq f(x) + \langle u, z - x \rangle,$$

for any $z \in C$. This condition is called the subgradient inequality.

According to the parametric representation of fuzzy number, The above subgradient inequality is equivalent to the following two inequalities:

$$f(z)_*(\alpha) \geq f(x)_*(\alpha) + \langle u_*(\alpha), z - x \rangle, \quad (3)$$

$$f(z)^*(\alpha) \geq f(x)^*(\alpha) + \langle u^*(\alpha), z - x \rangle. \quad (4)$$

for any $\alpha \in [0, 1]$. The set of all subgradients of f at x is called the subdifferential of f at x and is denoted by $\partial f(x)$. If $\partial f(x)$ is not empty, f is said to be subdifferential at x.

we now state an important theorem to show the relation between directional derivative of a fuzzy mapping and its subgradient under certain assumptions.

Theorem III.3. *Let C be a convex subset in $\mathbb{R}^n$ and $f : C \subset \mathbb{R}^n \to \mathcal{F}_0$ be a convex fuzzy mapping. Suppose that there exists the directional derivative of f at $x \in C$ in the direction $d \in \mathbb{R}^n$. If f is differentiable at $x \in \text{int(C)}$, then for any $d \in \mathbb{R}^n$*

$$f'(x; d) = \langle \nabla f(x), d \rangle. \quad (5)$$

Proof. It follows directly from the definition (III.1). □

Theorem III.4. *Let a fuzzy mapping $f : C \subset \mathbb{R}^n \to \mathcal{F}_0$ be differentiable at $x^0 \in \text{int(C)}$, then f is subdifferential at x^0, and $\nabla f(x^0) \in \partial f(x^0)$.*

Let f be a fuzzy mapping, and consider the problem to minimize $f(x)$ subject to $x \in S$. A point $x \in S$ is called a feasible solution to the problem. If $x^0 \in S$ and $f(x^0) \leq f(x)$ for each $x \in S$, then x^0 is called an optimal solution, or a global minimum point. If $x^0 \in S$ and there exists a neighborhood $N(x^0, \delta)$ around x^0 such that $f(x^0) \leq f(x)$ for each $x \in S \cap N(x^0, \delta)$, then x^0 is called a local optimal solution, or a local minimum point.

Theorem III.5. *(optimality conditions for the composite problem). Let $f : C \subset \mathbb{R}^n \to \mathcal{F}_0$ be a fuzzy mapping and $g : C \subset \mathbb{R}^n \to \mathcal{F}_0$ be a convex fuzzy mapping such that* $\mathrm{dom(g)} \subseteq \mathrm{int(dom(f))}$. *Consider the problem*

$$(P) \quad \min_{\mathbf{x}\in\mathbf{C}} \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x}).$$

(1) *(necessary condition) If $x^0 \in \mathrm{dom(g)}$ is a local optimal solution of (P) and f is differentiable at x^0, then*

$$-\nabla f(x^0) \in \partial g(x^0). \tag{6}$$

(2) *(necessary and sufficient condition for convex problems). Assume that f is convex fuzzy mapping. If f is differentiable at $x^0 \in \mathrm{dom(g)}$, then x^0 is a global optimal solution of (P) if and only if* (6) *is satisfied.*

Proof. It follow from the convexity of $\mathrm{dom(g)}$ that for any $x^0 = (x_1^0, x_2^0, ..., x_n^0), y = (y_1, y_2, ..., y_n) \in \mathrm{dom(g)}$ and $\lambda \in (0, 1)$, we have $z = (1-\lambda)x^0 + \lambda y \in \mathrm{dom(g)}$. using the local optimality of x^0, there exists a neighborhood $N(x^0, \delta)$ around x^0 such that

$$f(x^0) + g(x^0) \leq f(z) + g(z),$$

for any $z \in C \cap N(x^0, \delta)$.
That is,

$$f(x^0) + g(x^0) \leq f((1-\lambda)x^0 + \lambda y) + g((1-\lambda)x^0 + \lambda y).$$

Then

$$\begin{aligned} f(x^0)_*(\alpha) + g(x^0)_*(\alpha) \leq f((1-\lambda)x^0 + \lambda y)_*(\alpha) \\ + g((1-\lambda)x^0 + \lambda y)_*(\alpha), \end{aligned}$$

and

$$\begin{aligned} f(x^0)^*(\alpha) + g(x^0)^*(\alpha) \leq f((1-\lambda)x^0 + \lambda y)^*(\alpha) \\ + g((1-\lambda)x^0 + \lambda y)^*(\alpha). \end{aligned}$$

Using the convexity of g, it follows that

$$\begin{aligned} f(x^0)_*(\alpha) + g(x^0)_*(\alpha) \leq f((1-\lambda)x^0 + \lambda y)_*(\alpha) \\ + (1-\lambda)g(x^0)_*(\alpha) + \lambda g(y)_*(\alpha), \end{aligned}$$

and

$$\begin{aligned} f(x^0)^*(\alpha) + g(x^0)^*(\alpha) \leq f((1-\lambda)x^0 + \lambda y)^*(\alpha) \\ + (1-\lambda)g(x^0)^*(\alpha) + \lambda g(y)^*(\alpha), \end{aligned}$$

which is the same as

$$\frac{f((1-\lambda)x^0 + \lambda y)_*(\alpha) - f(x^0)_*(\alpha)}{\lambda} \geq g(x^0)_*(\alpha) - g(y)_*(\alpha),$$

$$\frac{f((1-\lambda)x^0 + \lambda y)^*(\alpha) - f(x^0)^*(\alpha)}{\lambda} \geq g(x^0)^*(\alpha) - g(y)^*(\alpha),$$

for any $\alpha \in [0, 1], \lambda \in (0, 1)$.
Taking the limit as $\lambda \to 0^+$ in the last inequality yields

$$f'(x^0; y - x^0)_*(\alpha) \geq g(x^0)_*(\alpha) - g(y)_*(\alpha), \tag{7}$$

$$f'(x^0; y - x^0)^*(\alpha) \geq g(x^0)^*(\alpha) - g(y)^*(\alpha), \tag{8}$$

since the directional derivatives of f at x^0 exist. In fact, by (5), we have

$$f'(x^0; y - x^0)_*(\alpha) = \big\langle \nabla f(x^0)_*(\alpha), y - x^0 \big\rangle,$$

$$f'(x^0; y - x^0)^*(\alpha) = \big\langle \nabla f(x^0)^*(\alpha), y - x^0 \big\rangle,$$

and hence utilizing (7) and (8), for any $y \in \mathrm{dom(g)}$, we get

$$g(y)_*(\alpha) \geq g(x^0)_*(\alpha) + \big\langle -\nabla f(x^0)_*(\alpha), y - x^0 \big\rangle,$$

$$g(y)^*(\alpha) \geq g(x^0)^*(\alpha) + \big\langle -\nabla f(x^0)^*(\alpha), y - x^0 \big\rangle,$$

Therefore

$$-\nabla f(x^0)_*(\alpha) \in \partial g(x^0)_*(\alpha),$$

$$-\nabla f(x^0)^*(\alpha) \in \partial g(x^0)^*(\alpha),$$

Thus theorem (II.7), implies that $-\nabla f(x^0) \in \partial g(x^0)$.

(b) If x^0 is an optimal solution of (P), then we already proved in part (a) that $-\nabla f(x^0) \in \partial g(x^0)$. Conversly, let that (6) is satisfied. therefore

$$-\nabla f(x^0)_*(\alpha) \in \partial g(x^0)_*(\alpha),$$

$$-\nabla f(x^0)^*(\alpha) \in \partial g(x^0)^*(\alpha),$$

for any $\alpha \in [0, 1]$. So,

$$g(y)_*(\alpha) \geq g(x^0)_*(\alpha) + \big\langle -\nabla f(x^0)_*(\alpha), y - x^0 \big\rangle, \tag{9}$$

$$g(y)^*(\alpha) \geq g(x^0)^*(\alpha) + \big\langle -\nabla f(x^0)^*(\alpha), y - x^0 \big\rangle, \tag{10}$$

for any $y \in \mathrm{dom(g)}$.
Utilizing the convexity of f, for any $y \in \mathrm{dom(g)}$, we have

$$f(y)_*(\alpha) \geq f(x^0)_*(\alpha) + \big\langle \nabla f(x^0)_*(\alpha), y - x^0 \big\rangle, \tag{11}$$

$$f(y)^*(\alpha) \geq f(x^0)^*(\alpha) + \big\langle \nabla f(x^0)^*(\alpha), y - x^0 \big\rangle. \tag{12}$$

Adding (9) and (11), also (10) and (12), we obtain

$$f(y)_*(\alpha) + g(y)_*(\alpha) \geq f(x^0)_*(\alpha) + g(x^0)_*(\alpha),$$

$$f(y)^*(\alpha) + g(y)^*(\alpha) \geq f(x^0)^*(\alpha) + g(x^0)^*(\alpha).$$

Hence

$$f(y) + g(y) \geq f(x^0) + g(x^0),$$

for any $y \in \mathrm{dom(g)}$. This means that x^0 is an optimal solution of (P). □

Application: These conclusions will play a key role in fuzzy optimization and will give rise the nonconvex composite model programming which will orient future research of the authors on fuzzy control problems.

REFERENCES

[1] L.A.Zadeh, "Fuzzy sets," information and control, vol. 8, pp. 338–353, 1965.

[2] S.S.L. Chang, L.A. Zadeh, On fuzzy mappings and control, IEEE Trans. Syst. Man Cybernet. 2 (1) (1972) 30–34.

[3] Roy Goetschel Jr., William Voxman, Elementary fuzzy calculus, Fuzzy Sets and Systems 18 (1986) 31–43.

[4] S. Seikkala, On the fuzzy initial value problem, Fuzzy Sets and Systems 24 (1987) 319–330.

[5] M.L. Puri, D.A. Ralescu, Differentials of fuzzy functions, Journal of Mathematical Analysis and Applications 91 (1983) 552–558.

[6] D. Dubois, H. Prade, Towards fuzzy differential calculus, Part 3: Differentiation, Fuzzy Sets and Systems 8 (1982) 225–233.

[7] D. Dubois, H. Prade, On several definitions of the differential of a fuzzy mapping, Fuzzy Sets and Systems 24 (1987) 117–120.

[8] M. Friedman, M. Ming, A. Kandel, Fuzzy derivatives and fuzzy cauchy problems using LP metric, in: Da Ruan (Ed.), Fuzzy Logic foundations and Applications, Kluwer, Dordrecht, 1996, pp. 57–72.

[9] James J. Buckley, Thomas Feuring, Fuzzy differential equations, Fuzzy Sets and Systems 110 (2000) 43–54.

[10] James J. Buckley, Thomas Feuring, Introduction to fuzzy partial differential equations, Fuzzy Sets and Systems 105 (1999) 247–248.

[11] S. Nanda, K. Kar, Convex fuzzy mappings, Fuzzy Sets and Systems 48 (1992) 129–132.

[12] Y.-R. Syau, On convex and concave fuzzy mappings, Fuzzy Sets and Systems 103 (1999) 163–168.

[13] D. Butnariu, Methods of solving optimization problems and linear equations in the space of fuzzy vectors, LibertasMath. 17 (1997) 1–7.

[14] M. Panigrahi, G. Panda, S. Nanda, Convex fuzzy mapping with differentiability and its application in fuzzy optimization, European J. Oper. Res., 185 (2008), 47–62.

[15] G.-X. Wang, C.-X. Wu, Directional derivatives and subdifferential of convex fuzzy mappings and application in convex fuzzy programming, Fuzzy Sets and Systems, 138 (2003), 559–591.

[16] P. Diamond, P. Kloeden, Metric Spaces of Fuzzy Sets, World Scientific, Singapore, 1994.

[17] A.Beck, First-Order Methods in Optimization, Siam(Society for Industrial and Applied Mathematics Philadelphia), 2018.

[18] C. Zhang, X.-H. Yuan, E. S. Lee, Duality theory in fuzzy mathematical programming problems with fuzzy coefficients, Comput. Math. Appl., 49 (2005), 1709–1730.

[19] B. Bede, Mathematics of Fuzzy Sets and Fuzzy Logic, Studies in Fuzziness and Soft Computing, vol.295, Springer, Heidelberg, 2013.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Modified Relay Node Placement in dense 3D Underwater Acoustic Sensor Networks

Zahra Mohammadi

Department of Electrical Engineering, Shahid Bahonar University of Kerman, Kerman, Iran, zahramohammadi@eng.uk.ac.ir

***Abstract*— The internet of underwater things (IoUTs) is regarded as one of the most technologies in smart oceans that Underwater acoustic sensor networks (UASNs) have recently emerged to support the concept of IoUT. Network lifetime is a vital requirement in UASNs because underwater sensors are supplied with batteries that are difficult to recharge. Relay node placement (RNP) is investigated in recent years to increase the network lifetime that the relay node plays an intermediate node between the critical node and one of its neighbours (i.e., line segment RNP). Due to the tiny feasible region in the line segment RNP, avoiding the redundant relay nodes is a challenging task, especially in the dense UASNs that the distance between nodes is short. In this paper, we design an efficient RNP by defining a threshold distance d_ϵ. Our objective is to make good tradeoffs between the performance of 3D UASN and complexity in relay- assistance networks. To do this, we introduce a modified difference convex approach (MoDCA). To evaluate our proposed method, we compare it with previous DCA and RIA in terms of the averag number of relay nodes (NoR_a) and network lifetime through two simulation scenarios.**

***Index Terms*—Underwater acoustic sensor network, Relay node, Network lifetime, Difference convex approach, dense deployment.**

I. Introduction

Underwater acoustic sensor networks (UASNs) can be used to explore precious resources such as oil and gas and monitor the underwater pipeline and cables. One of the main issues in the UASNs, which is vital to manage, is the network lifetime. Some solutions such as the routing protocol, relay node placement (RNP), and mobile-based data gathering have been recently studied to manage the energy consumption in the UASNs.

The design of an efficient routing protocol that balances energy consumption has been addressed in some existing works, such as [1] and [2]. In [1], to justify the traffic loads and to avoid the fast energy exhaustion of gateway nodes, the residual energy is considered as the data forwarding criterion. Based on the study in [2], besides the residual energy, the link quality is considered for routing cost and the data forwarding is based on path cost and depth of sensors. Based on the method in [3], the sensor nodes are classified into some clusters, and the cluster heads (CHs) are responsible for collecting data. Also, to prevent the hot region problem, the residual energy of sensors are taken into account to select the optimal CHs. Nothing that a hot region is created when some nodes consume a more considerable amount of energy compared with the other nodes. In [4], a method named clustered-based energy-efficient routing protocol has been proposed. This method has defined seven layers from top to seabed for the underwater environment by deploying powerful static courier nodes on different layers. Indeed, courier nodes in the bottom layer collect the information from sensor nodes and forward it to the base station (BS) by maximum power levels through courier nodes deployed in different layers. Overall, to review some of the routing protocols, we refer to [5] and the references therein.

Employing the autonomous underwater vehicle (AUV) to gather data from sensor nodes is an effective solution to reduce energy consumption in UASNs. For example, in [6], [7], by employing an AUV, data is collected from gateways and gateways are rotated over time to balance energy consumption. However, the transmission delays by AUV are very long that in the time-sensitive applications such as temperature and salinity evaluations for red tide forecast, it is necessary to collect data as soon as possible [8]. Furthermore, in AUV-aided underwater sensor networks the communication energy consumption is reduced through the reduction of network links, and therefore the connectivity cannot be guaranteed [6].

Furthermore, due to relay nodes' inability to generate data, they are employed to decrease the communication distance and prolong the network lifetime and thus in existing literature, various RNP approaches have been studied. For instance, the authors in [9] explore a heuristic-based relay-node adjustment (RA) scheme for placing relay nodes in the 3-dimensional UASN. In this technique, the position of relay nodes is determined to extend the network lifetime that gives rise to suboptimal RNP and thus can be caused resource wastage. To tackle the disadvantage of the suboptimal heuristic RNP methods, we proposed a novel difference convex (DC)-based RNP in [10] that is the line segment RNP. It is worth men-

tioning that, in the dense deployed sensors application, such as monitoring, the efficient use of relay nodes is important. In fact, in the dense deployed sensors application due to the short distance between nodes, relay nodes can not significantly improve the network lifetime. To tackle this issue in this paper we propose a Modified Difference Convex Approach (MoDCA) for RNP to avoid placing the redundant relays in the network. Our proposed MoDCA includes the RNP as a difference convex problem and has less system implementation complexity than the DCA scheme.

II. Network Model and problem formulation

A 3D relay-assistance UASN consists of a number of sensors to perform collaborative monitoring and data collection tasks in the search field, relays to reduce the communication distance between sensors and a base station (BS) as the central component that is used to gather data from distributed nodes. The set of sensors and relays are denoted as S and R, respectively. Two nodes i and j have a communication range (R_c) and can communicate when their distance (d_{ij}) is less than it. We use f_{ij} to show the rate of data from node i to node j. It is worth mentioning that the rate array beetwen nodes must met the following conditions

1) For each sensor node the sum of outgoing flow rates is the sum of incoming flow rates and generation rate.
2) For each relay node the sum of outgoing flow rates should be equal to sum of incoming flow rates.
3) The bandwidth of the acoustic channel in our model is limited and corresponds to the link capacity which means the sum of outgoing rates of each node must be less than that of its link capacity [9].

The Thorp's model for the underwater acoustic channel [11] is adopted in this paper that provides the attenuation for the underwater acoustic channel. Based on the Thorp's model the attenuation can be computed as:

$$A_{ij} = A_0 \times d_{ij}^{q} \times \alpha(f)^{d_{ij}} \tag{1}$$

where A_0 is the normalizing constant and q is the spreading factor in the underwater environment ($q \in \{1, 2\}$ depending on the water depth). Also, $\alpha(f)$ denotes the absorption coefficient in the underwater environment. By considering the threshold distance d_t and water depth H_s, the energy consumption in transmission of one bit data from node i to node j (i.e., p_{ij}) can be given as:

$$p_{ij} = p_s + \begin{cases} \alpha(f)^{d_{ij}} \times d_{ij}^2 & d_{ij} < d_t \\ \alpha(f)^{d_{ij}} \times d_{ij}^4 & d_t \le d_{ij} \le RT. \end{cases} \tag{2}$$

where p_s denotes the consumed energy for processing data, moreover d_{ij} is the 3D Euclidean distance computed as:

$$d_{ij}^2 = (x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2, \tag{3}$$

with the (x_i, y_i, z_i) and (x_j, y_j, z_j) as the coordinates of nodes i and j, respectively. The network lifetime is defined as the following :

$$\tau = \min_{i \in S \cup R} (\tau_i), \tag{4}$$

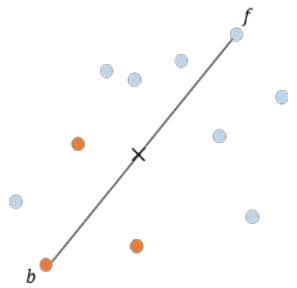


Fig. 1. Line segment RNP between critical node b and its neighbor f

with

$$\tau_i \times \Big(\sum_{j \in S \cup R \cup BS}^{i \neq j} p_{ij} f_{ij} + p_r \sum_{j \in S \cup R}^{k \neq i} f_{ki} \Big) = \epsilon_i, \tag{5}$$

that p_r and ϵ_i are the energy consumption for receiving one bit and the residual energy of the node i, respectively. Assuming the critical node of the network as b (i.e., the node with the minimum lifetime), based on the line segment RNP, the relay node is located on the line between b and one of its neighbors f as shown in Fig. 1. By considering the BS at the origin, the location vector of relay node (i.e., $\mathbf{r}$) in the line segment RNP is written as:

$$\mathbf{r} = \zeta \mathbf{b} + (1 - \zeta)\mathbf{f}, 0 \le \zeta \le 1 \tag{6}$$

that $\mathbf{b}$ and $\mathbf{f}$ show the location vector of node b and f, respectively. Followed by the DCA approach [10], we consider f is the farthest neighbor of b. By considering the energy consumption of relay node r and the lifetime of node c, the line segement RNP problem is given as the DC problem:

$$(\mathbf{r}, d_{br}, p_{br}, \zeta) = \arg\max\{\omega_1 g_\tau + (1 - \omega_1)(1 - g_p)\},$$

$$s.t. \tag{7a}$$

$$\tau_b = \frac{\epsilon_b}{(\sum_{j \neq f} f_{bj} \times p_{bj}) + (p_s \times \sum f_{kb}) + (f_{br} \times p_{br})}, \tag{7b}$$

$$p_{rf} = \begin{cases} p_s + \alpha(f)^{d_{rf}} \times (d_{rf})^4; d_t \le d_{rf} \le RT \\ p_s + \alpha(f)^{d_{rf}} \times (d_{rf})^2; \quad d_{rf} < d_t, \end{cases} \tag{7c}$$

$$p_{br} = \begin{cases} p_s + \alpha(f)^{d_{br}} \times d_{br}^4; \quad d_t \le d_{br} \le RT, \\ p_s + \alpha(f)^{d_{br}} \times d_{br}^2; \quad d_{br} < d_t, \end{cases} \tag{7d}$$

$$d_{br} = \|\mathbf{b} - \mathbf{r}\|, \tag{7e}$$

$$\mathbf{r} = \zeta \mathbf{b} + (1 - \zeta)\mathbf{f}, \tag{7f}$$

$$0 \le \zeta \le 1 \tag{7g}$$

that g_τ and g_p are the normalized b's lifetime and r's energy consumption. In what follows, we modify the line segment RNP for dense deployement of sensors.

III. Proposed Modified RNP strategy

As mentioned previously, relay nodes are employed to decrease the communication distance between nodes. The important problem in the sequential RNP is the dependency of relay deployment to each other. Indeed, in the multiple relay nodes placement, we aim to find the Cartesian coordinates of

$R = \{r_1, .., r_M\}$ to reduce the energy consumption of critical nodes $b_i, i = 1, .., M$ with lifetime of $\tau_{b_i}, i = 1, ..., M$ that:

$$\tau_{b_1} \leq \tau_{b_2} \leq ...\tau_{b_M} \tag{8}$$

By solving RNP, a set of optimal solutions $(x^*_{r_1}, ..., x^*_{r_M})$, $(y^*_{r_1}, ..., y^*_{r_M})$ and $(z^*_{r_1}, ..., z^*_{r_M})$ showing the Cartesian coordinates of R is obtained. These optimal solutions show the optimal value $[\tau^*_{b_1}, \tau^*_{b_2}, ...\tau^*_{b_M}]^T$. It is worth mentioning that in the efficient RNP, we have

$$\tau^*_{b_i} > \tau_{b_i}, i = 1, .., M \tag{9}$$

and when the above condition is not satisfied, the RNP must be stoped. In fact, in this situation the relay node r_i can not reduce the transmission distance of critical node. Here are two important things we need to know about RNP:

1) Due to the long transmission distance in the the sparce deployment of sensors relay nodes extend the network lifetime significantly.
2) Due to the short transmission distance in the the dense deployment of sensors increasing the relay nodes extend the network lifetime insignificantly.

This issue is shown in Fig. 2, means, the RNP problem is the scale-based or distance-based problem, where an efficient RNP is when the following condition satisfied:

$$d_{br} \geq d_\epsilon \tag{10}$$

that d_ϵ is selected according to the desired complexity-network lifetime trade-off, that can be cast as a linear constraint in the line segment optimization problem. However, at the first step, a stack of relay nodes, noted as S_r, will be built whose elements are in the order of adding to the underwater sensor network. In step i (i.e., placement of relay node r_i) the line segment RNP is modified as:

$$(\mathbf{r}_i, d_{b_i r_i}, p_{b_i r_i}, \zeta) = \arg\max\{\omega_1 g_\tau + (1-\omega_1)(1-g_p)\}, \quad s.t. \tag{11a}$$

$$\tau_{b_i} = \frac{\epsilon_{b_i}}{(\sum_{j \neq f} f_{b_i j} \times p_{b_i j}) + (p_s \times \sum f_{k b_i}) + (f_{b_i r_i} \times p_{b_i r_i})}, \tag{11b}$$

$$p_{r_i f_i} = \begin{cases} p_s + \alpha(f)^{d_{r_i f_i}} \times (d_{r_i f_i})^4; d_t \leq d_{r_i f_i} \leq RT \\ p_s + \alpha(f)^{d_{r_i f_i}} \times (d_{r_i f_i})^2; \quad d_{r_i f_i} < d_t, \end{cases} \tag{11c}$$

$$p_{b_i r_i} = \begin{cases} p_s + \alpha(f)^{d_{b_i r_i}} \times d^4_{b_i r_i}; \quad d_t \leq d_{b_i r_i} \leq RT, \\ p_s + \alpha(f)^{d_{b_i r_i}} \times d^2_{b_i r_i}; \quad d_{b_i r_i} < d_t, \end{cases} \tag{11d}$$

$$d_{b_i r_i} = \|\mathbf{b}_i - \mathbf{r}_i\|, \tag{11e}$$

$$\mathbf{r}_i = \zeta \mathbf{b}_i + (1-\zeta)\mathbf{f}_i, \tag{11f}$$

$$0 \leq \zeta \leq 1, \tag{11g}$$

$$d_{b_i r_i} \geq d_\epsilon. \tag{11h}$$

The difference convex algorithm algorithm tries to find a proper Cartesian coordinate for r_i that can be used to relay the traffic between (b_i, f_i). If the problem is successfully solved (i.e., the optimization problem is feasible), the network is ubdated and the algorithm tries to deploy subsequent relay node, otherwise the algorithm is stoped. In fact, in each step, a relay node must be introduced to increase the network lifetime. The pseudo-code of proposed MoDCA is given in algorithm 1.

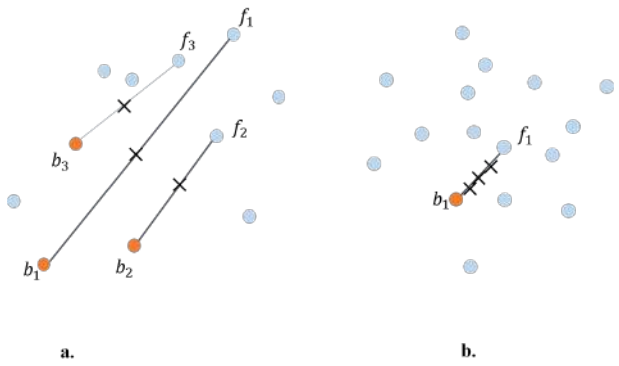


Fig. 2. RNP in two system scenario, a. sparce b. dense deployment

In what follows, we evaluate the performance of the proposed method through simulation operation.

Algorithm 1: Proposed MoDCA

Data: The set of all nodes (S), d_ϵ
Result: The adjusted relay nodes position $(x_i, y_i, z_i), \forall i$

```
Construct the stack S_r
R={}
Adjusted ⇐ False
while (S_r is not empty) and (Ajusted = False) do
    for each node n ∈ S ∪ R do
        Compute τ_n using 5
    end
    b_i ⇐ arg min τ_n, n ∈ S ∪ R
    r_i ⇐ ( POP(S_r);
    Define the problem 11
    Apply the difference convex algorithm to solve 11
    if the problem is infeasible
        Ajusted = True
    elseif
        R = R ∪ r_i
    endif
Return (x_i, y_i, z_i), ∀i, |R|
```

IV. Simulation Result and Comparison

In this section the proposed MoDCA scheme is evaluated and compared with DCA and RIA in terms of network lifetime and number of employed relays. The numeric parameter is selected as in [10], e.g, generation rate of each sensor node is set randomly from 10 to 200bit/Sec [9]. The initial energy of nodes is set to $4 \times 10^5 J$ and frequency of acoustic signal is 1kHz. Let us define the average number of relays (NoR_a) as:

$$NoR_a = \frac{\sum_{i=1}^{N} |R|_i}{N} \tag{12}$$

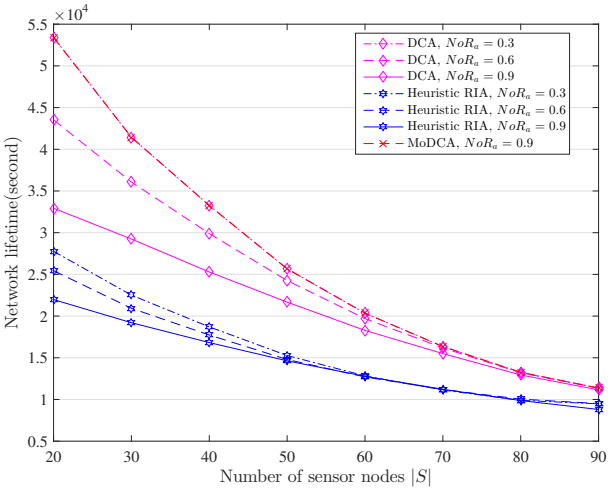


Fig. 3. Network lifetime versus number of sensors when $NoR_a = 0.9$

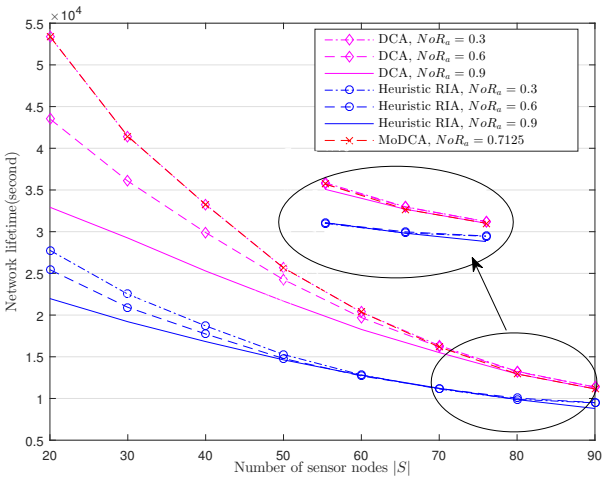


Fig. 4. Network lifetime versus number of sensors when $NoR_a = 0.7125$

that $|R|_i$ is the number of deployed relays when the number of sensors is i. Here we evaluate the network lifetime versus the number of sensors in Figs. 3 and 4.

As shown in Fig. 3 and 4, by increasing the number of sensor nodes the network lifetime is decreased. The reason for this is that at higher number of sensor nodes larger amounts of information is relayed and, therefore, greater energy is consumed. Moreover, in these scenarios, we evaluated the effect of increasing the number of relay nodes on the network lifetime. To do this, we change the NoR_a from 0.3 to 0.9. In the first scenario by considering $d_\epsilon = 0$, we reach to $NoR_a = 0.9$, in our proposed method. Given the results on Fig. 3, the proposed method has a same performance with DCA. As can be seen, although in all methods by increasing the number of relay nodes the network lifetime is increased, they reach to the saturation point. It is quite apparent that the use of extra relay nodes in this way cannot extend the lifetime of sensor nodes and, therefore, does not affect on the overall network lifetime significantly. This is because of the tiny feasible region to employ relays in the line segemnt RNP.

In the other evaluation, by increasing the d_ϵ in our proposed MoDCA, the $NoR_a = 0.7125$ is obtained in our method. The result of this experiment is shown in Fig. 4. As can be seen, in the our MoDCA proposed method by decreasing the number of employed relays provides a desirable network lifetime, means MoDCA avoids to deploy redundant relays in the network. It is quite apparent that the use of extra relay nodes in this way cannot extend the lifetime of sensor nodes and, therefore, does not affect on the overall network lifetime significantly.

V. Conclusion

The aim of this paper was to study RNP in dense 3D underwater acoustic sensor networks. To prevent the resource wastage, we introduce a threshold for line segement RNP that cast as a linear costraint in the difference convex- based RNP that we called it MoDCA. The simulation result shows our proposed MoDCA has a good tradeoffs between performance and averaged number of relays.

References

[1] S. M. Mazinani, H. Yousefi, and M. Mirzaie, "A vector-based routing protocol in underwater wireless sensor networks," *Wireless Personal Communications*, vol. 100, no. 4, pp. 1569–1583, 2018.

[2] A. Khasawneh, M. S. B. A. Latiff, O. Kaiwartya, and H. Chizari, "A reliable energy-efficient pressure-based routing protocol for underwater wireless sensor network," *Wireless Networks*, vol. 24, no. 6, pp. 2061–2075, 2018.

[3] K. Wang, H. Gao, X. Xu, J. Jiang, and D. Yue, "An energy-efficient reliable data transmission scheme for complex environmental monitoring in underwater acoustic sensor networks," *IEEE Sensors Journal*, vol. 16, no. 11, pp. 4051–4062, 2016.

[4] M. Ahmed, M. Salleh, and M. I. Channa, "Cbe2r: clustered-based energy efficient routing protocol for underwater wireless sensor network," *International Journal of Electronics*, vol. 105, no. 11, pp. 1916–1930, 2018.

[5] T. Islam and Y. K. Lee, "A comprehensive survey of recent routing protocols for underwater acoustic sensor networks," *Sensors*, vol. 19, no. 19, p. 4256, 2019.

[6] X. Zhuo, M. Liu, Y. Wei, G. Yu, F. Qu, and R. Sun, "Auv-aided energy-efficient data collection in underwater acoustic sensor networks," *IEEE Internet of Things Journal*, 2020.

[7] J. Yan, X. Yang, X. Luo, and C. Chen, "Energy-efficient data collection over auv-assisted underwater acoustic sensor network," *IEEE Systems Journal*, vol. 12, no. 4, pp. 3519–3530, 2018.

[8] N. Javaid, N. Ilyas, A. Ahmad, N. Alrajeh, U. Qasim, Z. A. Khan, T. Liaqat, and M. I. Khan, "An efficient data-gathering routing protocol for underwater wireless sensor networks," *Sensors*, vol. 15, no. 11, pp. 29 149–29 181, 2015.

[9] L. Liu, M. Ma, C. Liu, and Y. Shu, "Optimal relay node placement and flow allocation in underwater acoustic sensor networks," *IEEE Transactions on Communications*, vol. 65, no. 5, pp. 2141–2152, 2017.

[10] Z. Mohammadi, M. Soleimanpour-Moghadam, M. Askarizadeh, and S. Talebi, “Increasing the lifetime of underwater acoustic sensor networks: Difference convex approach,” *IEEE Systems Journal*, 2020.

[11] M. Stojanovic, “On the relationship between capacity and distance in an underwater acoustic communication channel,” *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 11, no. 4, pp. 34–43, 2007.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Multiobjective Zero-Sum Games in Intuitionistic fuzzy Environment

1st Hamid Bigdeli, 2nd Javad Tayyebi

1st Institute for the Study of War, Army Command and Staff University, Tehran, I.R. Iran, H.bigdeli@casu.ac.ir
2rd Department of Industrial Engineering, Birjand University of Technology, Birjand, I.R. Iran, Javadtayyebi@birjandut.ac.ir

***Abstract*— The aim of this paper is to develop a mathodology for solving a multiobjective zero-sum game with fuzzy payoffs. It is assumed that the payoffs of game matrices are expressed with triangular intuitionistic fuzzy numbers. The two programming problems are introduced to obtain the efficient strategies of the players. These problems are obtained by the concept of cuts of intuitionistic fuzzy sets. Finally, the validity and applicability of the proposed method is illustrated with a numerical example.**

***Index Terms*— Zero-sum game, Intutionistic fuzzy sets, Multiobjective optimization, Game theory, Fuzzy numbers.**

I. Introduction

The first research in fuzzy game theory dates back to the mid-1950s. Studies of fuzzy games have been made by incorporating fuzzy set theory. Fuzzy set theory plays an important role in modeling information uncertainty due to ambiguity and inaccuracy. Extensive researches have been done on the fuzzy games (for example see [2],[3], [4]). In real world affairs, decision makers have multiple goals and want to attain simultaneous goals. Therefore, it seems natural to use game theory in situations of conflict to control several goals. In this field, several studies have been made (for example see [2], [3]).

In 1983, Atanassov [1] introduced the concept of an intuitionistic fuzzy (IF) set. The intuitionistic fuzzy set which is characterized by two functions expressing the degree of membership and the degree of non-membership, respectively. For ambiguous information modeling, intuitionistic fuzzy sets provide more complete information than conventional fuzzy sets. It is seemed that intutionistic fuzzy sets are suitably express as an important factor which should be taken into account in players's hesitation degree. Li and Non [6] presented an interval-valued programming method to solve the matrix games with intuitionistic fuzzy payoffs. They proposed the linear programming problems to solve these games. Siekh etal. [7] considered the bimatrix games in intuitionistic environment. They presented the applications of the fuzzy intuitionistic numbers in the bimatrix games.

In this paper, we consider fuzzy multiobjective security games in which players' payoffs are expressed as intuitionistic fuzzy numbers and players' pure and mixed strategies are crisp. This problem has not been considered in previouse researches, based on the best knowledge of the authors.

The remainder of the paper is organized as follows. In section 2, some preliminaries, containing necessary notations and definitions of intuitionistic fuzzy sets, interval arithmetic and zero-sum games are presented. In section 3, a method is proposed for computing fuzzy values of two person zero-sum multiobjective games with fuzzy payoffs. In section 4, a numerical example is presented to apply the proposed method. Finally, conclusion is made in section 5.

II. Preliminaries

In this section, we provide some definitions and preliminaries of intuitionistic fuzzy sets and non-cooperative games according to [6] and [7].

Definition II.1. *Let $X = \{x_1, \cdots, x_n\}$ be a finite universal set. An intuitionistic fuzzy set $\tilde{A}$ in X is mathematically expressed as $\tilde{A} = \{\langle x_i, \mu_{\tilde{A}}(x_i), \vartheta_{\tilde{A}}(x_i)\rangle \mid x_i \in X\}$ in which $\mu_{\tilde{A}} : X \to [0,1]$ and $\vartheta_{\tilde{A}} : X \to [0,1]$ are respectively the membership degree and the non-membership degree of an element $x_i \in X$ to the set $\tilde{A}$ such that they satisfy the inequality $\mu_{\tilde{A}}(x_i) + \vartheta_{\tilde{A}}(x_i) \leq 1$ for all $x_i \in X$.*

Definition II.2. *$((\alpha, \beta) - cuts)$, Let $\alpha, \beta \in [0,1]$ be fixed numbers such that $\alpha + \beta \leq 1$. An (α, β)-cut generated by an intuitionistic fuzzy set $\tilde{A}$ is defined as*

$$\tilde{A}_{\alpha,\beta} = \{x \mid \mu_{\tilde{A}}(x) \geq \alpha, \vartheta_{\tilde{A}}(x) \leq \beta\}.$$

So $\tilde{A}_{\alpha,\beta}$ is a crisp set of elements $x \in X$, which belong to $\tilde{A}$ at least with the degree α and which do not belong to $\tilde{A}$ at most with the degree β.

In the following, a special type of intuitionistic fuzzy sets is introduced.

Definition II.3. *A triangular intuitionistic fuzzy number (TIFN)* $\tilde{a} = (a, l_{\tilde{a}}, r_{\tilde{a}}; w_{\tilde{a}}, u_{\tilde{a}})$ *is a special intuitionistic fuzzy number, whose membership and non-membership functions are defined as follows:*

$$\mu_{\tilde{a}}(x) = \begin{cases} \dfrac{x - a + l_{\tilde{a}}}{l_{\tilde{a}}} w_{\tilde{a}} & a - l_{\tilde{a}} \le x < a, \\ \dfrac{a + r_{\tilde{a}} - x}{r_{\tilde{a}}} w_{\tilde{a}} & a \le x < a + r_{\tilde{a}}, \\ 0 & \textit{otherwise}, \end{cases}$$

and

$$\vartheta_{\tilde{a}}(x) = \begin{cases} \dfrac{(a - x) + u_{\tilde{a}}(x - a + l_{\tilde{a}})}{l_{\tilde{a}}} & a - l_{\tilde{a}} \le x < a, \\ \dfrac{(x - a) + u_{\tilde{a}}(a + r_{\tilde{a}} - x)}{r_{\tilde{a}}} & a \le x < a + r_{\tilde{a}}, \\ 1 & \textit{otherwise}, \end{cases}$$

where $l_{\tilde{a}}$, $r_{\tilde{a}}$ *are respectively the left and right spreads and* a *is the mean value.* $w_{\tilde{a}}$ *and* $u_{\tilde{a}}$ *represent the maximum degree of membership and minimum degree of non-membership, respectively. Moreover, they satisfy the following conditions*

$$w_{\tilde{a}}, u_{\tilde{a}} \ge 0, \quad w_{\tilde{a}} + u_{\tilde{a}} \le 1.$$

Definition II.4. *A* (α, β)*-cut set of a triangular intuitionistic fuzz number* $\tilde{a} = (a, l_{\tilde{a}}, r_{\tilde{a}}; w_{\tilde{a}}, u_{\tilde{a}})$ *is a crisp subset of* $\mathbb{R}$ *which is defined as* $\tilde{a}_{\alpha,\beta} = \{x \mid \mu_{\tilde{A}}(x) \ge \alpha, \vartheta_{\tilde{A}}(x) \le \beta\}$ *where* $0 \le \alpha \le w_{\tilde{a}}$, $u_{\tilde{a}} \le \beta \le 1$ *and* $0 \le \alpha + \beta \le 1$.

Definition II.5. *A* α*-cut set of a triangular intuitionistic fuzzy numbers* $\tilde{a} = (a, l_{\tilde{a}}, r_{\tilde{a}}; w_{\tilde{a}}, u_{\tilde{a}})$ *is a crisp subset of* $\mathbb{R}$ *which is defined as* $\tilde{a}_\alpha = \{x \mid \mu_{\tilde{A}}(x) \ge \alpha\}$ *where* $0 \le \alpha \le w_{\tilde{a}}$.

for any $\alpha \in [0, w_{\tilde{a}}]$ easily we can see that $\tilde{a}_\alpha = [l_{\tilde{a}} + \dfrac{\alpha}{w_{\tilde{a}}}(a - l_{\tilde{a}}), a - \dfrac{\alpha}{w_{\tilde{a}}}(r_{\tilde{a}} - a)]$.

Definition II.6. *A* β*-cut set of a triangular intuitionistic fuzzy numbers* $\tilde{a} = (a, l_{\tilde{a}}, r_{\tilde{a}}; w_{\tilde{a}}, u_{\tilde{a}})$ *is a crisp subset of* $\mathbb{R}$ *which is defined as* $\tilde{a}_\beta = \{x \mid \vartheta_{\tilde{A}}(x) \le \beta\}$ *where* $u_{\tilde{a}} \le \beta \le 1$.

for any $\beta \in [u_{\tilde{a}}, 1]$ easily we can see that $\tilde{a}_\beta = [\dfrac{[(1-\beta)\tilde{a} + (\beta - u_{\tilde{a}})l_{\tilde{a}}]}{1 - u_{\tilde{a}}}, \dfrac{[(1-\beta)\tilde{a} + (\beta - u_{\tilde{a}})r_{\tilde{a}}]}{1 - u_{\tilde{a}}}]$.

Ranking triangular intuitionistic fuzzy numbers is a difficult issue. In this paper, a new ranking order relation of triangular intuitionistic fuzzy numbers is defined as in the following Definition.

Definition II.7. *Let* $\tilde{a} = (a, l_{\tilde{a}}, r_{\tilde{a}}; w_{\tilde{a}}, u_{\tilde{a}})$ *and* $\tilde{b} = (b, l_{\tilde{b}}, r_{\tilde{b}}; w_{\tilde{b}}, u_{\tilde{b}})$ *be two triangular intuitionistic fuzzy numbers.* $\tilde{a}_\alpha$ *and* $\tilde{b}_\alpha$ *are any* α*-cut sets of* $\tilde{a}$ *and* $\tilde{b}$. $\tilde{a}_\beta$ *and* $\tilde{b}_\beta$ *are any* β*-cut sets of* $\tilde{a}$ *and* $\tilde{b}$*, respectively. Then, we stipulate the following relations*
1) $\tilde{a} \tilde{\le} \tilde{b}$ *if and only if* $\tilde{a}_\alpha \le \tilde{b}_\alpha$ *and* $\tilde{a}_\beta \le \tilde{b}_\beta$ *for any given* $\alpha \in [0, min\{w_{\tilde{a}}, w_{\tilde{b}}\}]$ *and* $\beta \in [max\{u_{\tilde{a}}, u_{\tilde{b}}\}, 1]$*, where* $0 \le \alpha + \beta \le 1$*;*
2) $\tilde{a} \tilde{\ge} \tilde{b}$ *if and only if* $\tilde{a}_\alpha \ge \tilde{b}_\alpha$ *and* $\tilde{a}_\beta \ge \tilde{b}_\beta$ *for any given* $\alpha \in [0, min\{w_{\tilde{a}}, w_{\tilde{b}}\}]$ *and* $\beta \in [max\{u_{\tilde{a}}, u_{\tilde{b}}\}, 1]$*, where* $0 \le \alpha + \beta \le 1$.

The symbols $\tilde{\le}$ and $\tilde{\ge}$ represent "approximately less than or equal to" and "approximately greater than or equal to", respectively.

A. Zero-sum game

Let two players of a two-person zero-sum game be denoted by Players I and II. Assume that $I = \{1, \ldots, m\}$ and $J = \{1, \ldots, n\}$ are the sets of pure strategies of Players I and II, respectively. When Player I chooses the pure strategy i and Player II chooses the pure strategy j, then a_{ij} is the payoff for Player I and $-a_{ij}$ is the payoff for Player II. The two-person zero-sum matrix game G in normal form can be represented as a payoff matrix $A = [a_{ij}]_{m \times n}$. Consider the game G with no saddle point, i.e. $\max_i \min_j a_{ij} \ne \min_j \max_i a_{ij}$. To solve such a game, the concept of mixed strategy for Players I and II is introduced. Mixed strategy spaces are denoted for Players I and II as follows, respectively:

$$X = \{x \in \mathbb{R}^m | \sum_{i=1}^{m} x_i = 1, x_i \ge 0, i = 1, \ldots, m\}, \tag{3}$$

$$Y = \{y \in \mathbb{R}^n | \sum_{j=1}^{n} y_j = 1, y_j \ge 0, j = 1, \ldots, n\}. \tag{4}$$

In fact the mixed strategies for Players I and II are the probability distributions on the sets I and J, respectively. It is conventional to assume that Player I is a maximizing player while Player II is a minimizing player. Further, for $x \in X$, $y \in Y$, the scalar $x^T A y$ is the expected payoff to Player I, and as the game is zero sum, the expected payoff to Player II is $-x^T A y$. We know that for a two-person zero-sum game G with payoff matrix A, we have

$$\max_{x \in X} \min_{y \in Y} x^T A y = \min_{y \in Y} \max_{x \in X} x^T A y.$$

A pair of strategies (x^*, y^*) satisfying the above equation is called an equilibrium solution [2].

III. Multiobjective two-person zero-sum game with intuitionistic fuzzy payoffs

In this section, we consider a multiobjective game problem in intuitionistic fuzzy environment. Assume that each player has p objectives. The following multiple fuzzy payoff matrices represent a multiobjective two person zero-sum game with intuitionistic fuzzy payoffs:

$$\tilde{A}^1 = \begin{bmatrix} \tilde{a}^1_{11} & \ldots & \tilde{a}^1_{1n} \\ \vdots & \ddots & \vdots \\ \tilde{a}^1_{m1} & \ldots & \tilde{a}^1_{mn} \end{bmatrix}, \quad \ldots \quad , \tilde{A}^p = \begin{bmatrix} \tilde{a}^p_{11} & \ldots & \tilde{a}^p_{1n} \\ \vdots & \ddots & \vdots \\ \tilde{a}^p_{m1} & \ldots & \tilde{a}^p_{mn} \end{bmatrix}$$

where $\tilde{A}^k$ is the payoff matrix of the game with respect to the k-th objective function, for $k = 1, \ldots, p$. Mixed strategy

spaces are denoted for Players I and II as (3) and (4), respectively.

Choosing $x \in X$ and $y \in Y$ by Players I and II, respectively, we imply that the expected payoff of the game is

$$\tilde{v}(x,y) = x^T \tilde{A} y = [\tilde{v}^1(x,y), \ldots, \tilde{v}^p(x,y)] \tag{8}$$

where $\tilde{A} = [\tilde{A}^1, \ldots, \tilde{A}^p]$ and $\tilde{v}^k(x,y) = x^T \tilde{A}^k y$, $k = 1, \ldots, p$.

We can obtain player I's maximin efficient strategy $x^* \in X$ through solving the following problem:

$$\begin{aligned} &\max \left(\min_{y \in Y} \tilde{v}^1(x,y), \ldots, \min_{y \in Y} \tilde{v}^p(x,y)\right) \\ &s.t. \quad \sum_{i=1}^{m} x_i = 1 \\ &\qquad x_i \geq 0 \quad i = 1, \ldots, m. \end{aligned} \tag{14}$$

We set,

$$\underline{\tilde{v}}^k = \min_{y \in Y} \tilde{v}^k(x,y) \quad k = 1, \ldots, p, \tag{9}$$

Thus, we have

$$\begin{aligned} &\max (\underline{\tilde{v}}^1, \ldots, \underline{\tilde{v}}^p) \\ &s.t. \\ &\qquad \sum_{j=1}^{n} \sum_{i=1}^{m} x_i \tilde{a}_{ij}^1 y_j \tilde{\geq} \underline{\tilde{v}}^1 \quad \forall y \in Y \\ &\qquad \vdots \\ &\qquad \sum_{j=1}^{n} \sum_{i=1}^{m} x_i \tilde{a}_{ij}^p y_j \tilde{\geq} \underline{\tilde{v}}^p \quad \forall y \in Y \\ &\qquad \sum_{i=1}^{m} x_i = 1 \\ &\qquad x_i \geq 0 \quad i = 1, \ldots, m. \end{aligned} \tag{15}$$

Since "$\tilde{\geq}$" and "$\tilde{\leq}$" preserve the linear ranking order relations, we consider only the extreme points of the sets X. Hence,

$$\begin{aligned} &\max \ (\underline{\tilde{v}}^1, \ldots, \underline{\tilde{v}}^p) \\ &s.t. \\ &\qquad \sum_{i=1}^{m} \tilde{a}_{ij}^1 x_i \tilde{\geq} \underline{\tilde{v}}^1 \quad j = 1, \ldots, n \\ &\qquad \vdots \\ &\qquad \sum_{i=1}^{m} \tilde{a}_{ij}^p x_i \tilde{\geq} \underline{\tilde{v}}^p \quad j = 1, \ldots, n \\ &\qquad \sum_{i=1}^{m} x_i = 1 \\ &\qquad x_i \geq 0 \quad i = 1, \ldots, m. \end{aligned} \tag{16}$$

On the other hand, we can represent (α, β)-cuts of the triangular intuitionistic fuzzy numbers as $\underline{\tilde{v}}^k = \{(\underline{v}^k)_\alpha, (\underline{v}^k)_\beta\} = \{[(\underline{v}^{kL})_\alpha, (\underline{v}^{kR})_\alpha], [(\underline{v}^{kL})_\beta, (\underline{v}^{kR})_\beta]\}$.

Using (α, β)-cuts of the triangular intuitionistic fuzzy numbers and the weighted sum method in multiobjective optimization, we propose the following two problems to obtain the efficient optimal solutions for the players I and II, respectively. Thus,

$$\begin{aligned} &\max \ \sum_{k=1}^{p} \lambda_k ((\underline{v}^{kL})_\alpha + (\underline{v}^{kL})_\beta) \\ &s.t. \\ &\qquad \sum_{k=1}^{p} \lambda_k ((\underline{v}^{kL})_\alpha + (\underline{v}^{kL})_\beta) \\ &\qquad \leq \sum_{k=1}^{p} \lambda_k ((\underline{v}^{kR})_\alpha + (\underline{v}^{kR})_\beta) \\ &\qquad \sum_{i=1}^{m} (\tilde{a}_{ij}^{1L})_\alpha x_i \geq_I (\underline{v}^{1L})_\alpha \quad j = 1, \ldots, n \\ &\qquad \sum_{i=1}^{m} (\tilde{a}_{ij}^{1R})_\alpha x_i \geq_I (\underline{v}^{1R})_\alpha \quad j = 1, \ldots, n \\ &\qquad \sum_{i=1}^{m} (\tilde{a}_{ij}^{1L})_\beta x_i \geq_I (\underline{v}^{1L})_\beta \quad j = 1, \ldots, n \\ &\qquad \sum_{i=1}^{m} (\tilde{a}_{ij}^{1R})_\beta x_i \geq_I (\underline{v}^{1R})_\beta \quad j = 1, \ldots, n \\ &\qquad \vdots \\ &\qquad \sum_{i=1}^{m} (\tilde{a}_{ij}^{pL})_\alpha x_i \geq_I (\underline{v}^{pL})_\alpha \quad j = 1, \ldots, n \\ &\qquad \sum_{i=1}^{m} (\tilde{a}_{ij}^{pR})_\alpha x_i \geq_I (\underline{v}^{pR})_\alpha \quad j = 1, \ldots, n \\ &\qquad \sum_{i=1}^{m} (\tilde{a}_{ij}^{pL})_\beta x_i \geq_I (\underline{v}^{pL})_\beta \quad j = 1, \ldots, n \\ &\qquad \sum_{i=1}^{m} (\tilde{a}_{ij}^{pR})_\beta x_i \geq_I (\underline{v}^{pR})_\beta \quad j = 1, \ldots, n \\ &\qquad \sum_{i=1}^{m} x_i = 1 \\ &\qquad x_i \geq 0 \quad i = 1, \ldots, m \end{aligned} \tag{17}$$

where $\lambda \in \Lambda = \left\{ \lambda \in \mathbb{R}^p \,\middle|\, \lambda \geq 0, \sum_{k=1}^{p} \lambda_k = 1 \right\}$, and can be interpreted as the relative importance of the k-th objective function to Player I.

And for player II,

$$\begin{aligned} &\min \ \sum_{k=1}^{p} \lambda_k ((\bar{v}^{kL})_\alpha + (\bar{v}^{kL})_\beta) \\ &s.t. \\ &\qquad \sum_{k=1}^{p} \lambda_k ((\bar{v}^{kL})_\alpha + (\bar{v}^{kL})_\beta) \\ &\qquad \leq \sum_{k=1}^{p} \lambda_k ((\bar{v}^{kR})_\alpha + (\bar{v}^{kR})_\beta) \\ &\qquad \sum_{j=1}^{n} (\tilde{a}_{ij}^{1L})_\alpha y_j \geq_I (\bar{v}^{1L})_\alpha \quad i = 1, \ldots, m \\ &\qquad \sum_{j=1}^{n} (\tilde{a}_{ij}^{1R})_\alpha y_j \geq_I (\bar{v}^{1R})_\alpha \quad i = 1, \ldots, m \\ &\qquad \sum_{j=1}^{n} (\tilde{a}_{ij}^{1L})_\beta y_j \geq_I (\bar{v}^{1L})_\beta \quad i = 1, \ldots, m \\ &\qquad \sum_{j=1}^{n} (\tilde{a}_{ij}^{1R})_\beta y_j \geq_I (\bar{v}^{1R})_\beta \quad i = 1, \ldots, m \\ &\qquad \vdots \\ &\qquad \sum_{j=1}^{n} (\tilde{a}_{ij}^{pL})_\alpha y_j \geq_I (\bar{v}^{pL})_\alpha \quad i = 1, \ldots, m \\ &\qquad \sum_{j=1}^{n} (\tilde{a}_{ij}^{pR})_\alpha y_j \geq_I (\bar{v}^{pR})_\alpha \quad i = 1, \ldots, m \\ &\qquad \sum_{j=1}^{n} (\tilde{a}_{ij}^{pL})_\beta y_j \geq_I (\bar{v}^{pL})_\beta \quad i = 1, \ldots, m \\ &\qquad \sum_{j=1}^{n} (\tilde{a}_{ij}^{pR})_\beta y_j \geq_I (\bar{v}^{pR})_\beta \quad i = 1, \ldots, m \\ &\qquad \sum_{j=1}^{n} y_j = 1 \\ &\qquad y_j \geq 0 \quad j = 1, \ldots, n \end{aligned}$$

A. Numerical Example

Consider the payoff matrices of Players I and II as follows:

$$\tilde{A}_1 = \begin{bmatrix} <(175,180,190);0.6,0.2> & <(150,156,158);0.6,0.1> \\ <(80,90,100);0.9,0.1> & <(175,180,190);0.6,0.2> \end{bmatrix},$$

$$\tilde{A}_2 = \begin{bmatrix} <(125,130,135);0.8,0.2> & <(120,130,135);0.5,0.5> \\ <(120,130,135);0.6,0.3> & <(150,160,170);0.6,0.2> \end{bmatrix}.$$

This problem is a biobjective two person zero-sum game.
Assume that the importance of objective functions to Player I are the same. So we set $\lambda_1 = \lambda_2 = 0.5$.
The problem (PI-(α,β)) for these data is as follows:

$$\begin{aligned}
&\max \ \tfrac{1}{2}((\underline{v}^{1L})_\alpha + (\underline{v}^{1L})_\beta) + \tfrac{1}{2}((\underline{v}^{2L})_\alpha + (\underline{v}^{2L})_\beta) \\
&s.t. \\
&\quad \tfrac{1}{2}((\underline{v}^{1L})_\alpha + (\underline{v}^{1L})_\beta) + \tfrac{1}{2}((\underline{v}^{2L})_\alpha + (\underline{v}^{2L})_\beta) \leq \\
&\quad \tfrac{1}{2}((\underline{v}^{1R})_\alpha + (\underline{v}^{1R})_\beta) + \tfrac{1}{2}((\underline{v}^{2R})_\alpha + (\underline{v}^{2R})_\beta) \\
&\quad (175 + \tfrac{5\alpha}{0.6})x_1 + (80 + \tfrac{10\alpha}{0.9})x_2 \geq (\underline{v}^{1L})_\alpha \\
&\quad (150 + \tfrac{6\alpha}{0.6})x_1 + (175 + \tfrac{5\alpha}{0.6})x_2 \geq (\underline{v}^{1L})_\alpha \\
&\quad (180 - \tfrac{10\alpha}{0.6})x_1 + (90 - \tfrac{10\alpha}{0.9})x_2 \geq (\underline{v}^{1R})_\alpha \\
&\quad (156 - \tfrac{2\alpha}{0.6})x_1 + (180 - \tfrac{10\alpha}{0.6})x_2 \geq (\underline{v}^{1R})_\alpha \\
&\quad (125 + \tfrac{5\alpha}{0.8})x_1 + (120 + \tfrac{10\alpha}{0.6})x_2 \geq (\underline{v}^{2L})_\alpha \\
&\quad (120 + \tfrac{10\alpha}{0.5})x_1 + (150 + \tfrac{10\alpha}{0.6})x_2 \geq (\underline{v}^{2L})_\alpha \\
&\quad (130 - \tfrac{5\alpha}{0.8})x_1 + (130 - \tfrac{5\alpha}{0.6})x_2 \geq (\underline{v}^{2R})_\alpha \\
&\quad (130 - \tfrac{5\alpha}{0.5})x_1 + (130 - \tfrac{5\alpha}{0.6})x_2 \geq (\underline{v}^{2R})_\alpha \\
&\quad (\tfrac{180(1-\beta)+175(\beta-0.2)}{0.8})x_1 + (\tfrac{90(1-\beta)+80(\beta-0.1)}{0.9})x_2 \geq (\underline{v}^{1L})_\beta \\
&\quad (\tfrac{156(1-\beta)+150(\beta-0.1)}{0.9})x_1 + (\tfrac{180(1-\beta)+175(\beta-0.2)}{0.8})x_2 \geq (\underline{v}^{1L})_\beta \\
&\quad (\tfrac{180(1-\beta)+190(\beta-0.2)}{0.8})x_1 + (\tfrac{90(1-\beta)+100(\beta-0.1)}{0.9})x_2 \geq (\underline{v}^{1R})_\beta \\
&\quad (\tfrac{156(1-\beta)+158(\beta-0.1)}{0.9})x_1 + (\tfrac{180(1-\beta)+190(\beta-0.2)}{0.8})x_2 \geq (\underline{v}^{1R})_\beta \\
&\quad (\tfrac{130(1-\beta)+125(\beta-0.2)}{0.8})x_1 + (\tfrac{130(1-\beta)+120(\beta-0.3)}{0.7})x_2 \geq (\underline{v}^{2L})_\beta \\
&\quad (\tfrac{130(1-\beta)+120(\beta-0.5)}{0.8})x_1 + (\tfrac{130(1-\beta)+120(\beta-0.3)}{0.7})x_2 \geq (\underline{v}^{2L})_\beta \\
&\quad (\tfrac{130(1-\beta)+135(\beta-0.2)}{0.5})x_1 + (\tfrac{130(1-\beta)+135(\beta-0.3)}{0.7})x_2 \geq (\underline{v}^{2R})_\beta \\
&\quad (\tfrac{130(1-\beta)+135(\beta-0.5)}{0.5})x_1 + (\tfrac{160(1-\beta)+170(\beta-0.2)}{0.8})x_2 \geq (\underline{v}^{2R})_\beta \\
&\quad x_1 + x_2 = 1 \\
&\quad x_1, x_2 \geq 0
\end{aligned}$$

Considering the values of α and β, we can obtain the efficient optimal solutions. For example, for $\alpha = 0, \beta = 1$, we have
$<(\underline{v}^{1L})_\alpha, (\underline{v}^{1L})_\beta, (\underline{v}^{2L})_\alpha, (\underline{v}^{2L})_\beta, (\underline{v}^{1R})_\alpha, (\underline{v}^{1R})_\beta, (\underline{v}^{2R})_\alpha, (\underline{v}^{2R})_\beta> = < 155.20, 155.20, 123.95, 123.95, 0, 0, 0, 0 >$

$x^* = 0.79, y^* = 0.21$

Also, for Player II is similar.

The above problem have been solved by Lingo software.

IV. Conclusion

The zero-sum multiobjective game with fuzzy intuitionistic payoffs considered in this research. A method presented to find an efficient strategy and the objective values in such problems. The fuzzy payoffs are considered to be triangular fuzzy intuitionistic numbers. The problem is converted to a multiobjective game problem with interval payoffs by considering the concept of α-cuts and β-cuts, and its solutions are obtained by solving an interval multiobjective linear programming problem. Then, we proposed a mathematical programming problem with the crisp coefficients. Finally, a numerical example is proposed to compute an efficient strategy for Players. The proposed method can be applied to interval-valued matrix games simply.

References

[1] K. T. Atanassov, Intuitionistic fuzzy sets, In: V. Sgurev, ed., VII ITKR's session, Sofia, June 1983.

[2] H. Bigdeli, H. Hassanpour, *A satisfactory strategy of multi-objective two person matrix games with fuzzy payoffs*, Iranian Journal of Fuzzy Systems, 13, 17-33 (2016).

[3] H. Bigdeli, H. Hassanpour, An approach to solve multiobjective linear production planning games with fuzzy parameters,*Yugoslav Journal of Operations Research,* **1** *2(2018).*

*[4] H. Bigdeli, H. Hassanpour, J. Tayyebi, Multiobjective security games with fuzzy payoffs,*Iranian Journal of Fuzzy Systems, **16** 1, 89-101(2019).

[5] H. Bigdeli, J. Tayyebi, Cooperative Games with Multiple Scenarios in Intuitionistic Fuzzy Environment,*Caspian Journal of Mathematical Sciences (CJMS), Accepted(2021).*

[6] D. Li, J. Nan, 'An interval-valued programming approach to matrix games with payoffs of triangular intuitionistic fuzzy numbers', Iranian Journal of Fuzzy Systems, 11(2), pp. 45-57 (2014). doi: 10.22111/ijfs.2014.1490.

[7] Seikh, M.R., Pal, M. and Nayak, P.K., Application of triangular intuitionistic fuzzy numbers in bi-matrix games. International Journal of Pure and Applied Mathematics, **79(2)** , 235-247 (2012).

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

New results on prime A-ideals in MV-semimodules

1st Simin Saidi Goraghani and 2nd Rajab Ali Borzooei

1st Department of Mathematics, Farhangian University, Tehran, Iran, siminsaidi@yahoo.com
2nd Department of Mathematics, Shahid Beheshti University,Tehran,Iran, borzooei@sbu.ac.ir

***Abstract*— Recently, we introduced a new definition of MV-semimodules [8]. In this paper, we study prime A-ideals in MV-semimodules, and we obtain some results that could be proved by the old definitions only under certain conditions. Finally, we present some equivalent conditions for having prime A-ideals in MV-semimodules.**

***Index Terms*— MV-$semi$ring, MV-semimodule, Q-ideal, prime A-ideal.**

I. Introduction

The concept of semirings was introduced by H. S. Vandiver (1935). MV-semiring is an special ring that has recently been introduced. Nowadays, the theory of idempotent semirings has many applications in other fields. The concept of MV-algebras was defined by C.C. Chang as algebras corresponding to the Łukasiewicz Logic. In 2003, A. Di Nola presented the notion of MV-modules as a PMV-algebra multiplication in an MV-algebra. Then some researches worked on MV-modules (see, for instance, [5–7]), and some papers were written by that definition. In [2], Di Nola and C. Russo introduced the notion of MV-semimodules as an MV-semiring multiplication in an abelian monoid. In [8], we presented the notion of *new definition of MV-semimodules*. We considered multiplication of an MV-semiring in an MV-algebra in order to verify MV-modules and their ideals better and simpler than before.
In this paper, we study prime A-ideals in MV-semimodules, and we obtain some results that could be proved by the old definitions only under certain conditions. For example, we present some equivalent conditions for having prime A-ideals in MV-semimodules.

II. Preliminaries

The first, we review the definitions that we need in the following sections.

Definition 1. *[2] An algebraic structure $(\Upsilon, \dot{+}, ., 0, 1)$ of type $(2, 2, 0, 0)$ is called a semiring if;*
(i) $(\Upsilon, \dot{+}, 0)$ is an abelian monoid,
(ii) $(\Upsilon, ., 1)$ is a monoid,
(iii) "$.$" distributes over "$\dot{+}$" from either side. Υ is said commutative if $t.h = h.t$, for every $t, h \in \Upsilon$. If Υ satisfies the equation $t \dot{+} t = t$, for every $t \in \Upsilon$, then Υ is called idempotent.

Consider $\emptyset \neq \Omega \subseteq \Upsilon$. Ω is called a left ideal of semiring Υ if it satisfies the following condition:
(1) if $t, h \in \Omega$, then $t \dot{+} h \in \Omega$,
(2) if $t \in \Omega$ and $\mu \in \Upsilon$, then $\mu.t \in \Omega$.

Semiring $(\Upsilon, \dot{+}, ., 0, 1)$ is called An MV-semiring, if it is commutative and additive idempotent. Moreover, there exists a map $\diamond : \Upsilon \to \Upsilon$ that satisfying the following conditions:
(i) $\mu.\nu = 0$ if and only if $\nu \leq \mu^{\diamond}$ (" $\leq$ " is naturally defined by means of $\dot{+}$),
(ii) $\mu + \nu = (\mu^{\diamond}.(\mu^{\diamond}.\nu)^{\diamond})^{\diamond}$,
for every $\mu, \nu \in \Upsilon$.

Definition 2. *[1] Algebra $V = (V, \oplus, \diamond, 0)$ of type $(2, 1, 0)$ is called an MV-algebra, if it satisfying the following equations*
$(MV1)$ $(V, \oplus, 0)$ is an Abelian monoid,
$(MV2)$ $(\mu^{\diamond})^{\diamond} = \mu$,
$(MV3)$ $0^{\diamond} \oplus \mu = 0^{\diamond}$,
$(MV4)$ $(\mu^{\diamond} \oplus \nu)^{\diamond} \oplus \nu = (\nu^{\diamond} \oplus \mu)^{\diamond} \oplus \mu$, for every $\mu, \nu \in V$.
If we define the constant $1 = 0^{\diamond}$ and operations $\odot$ and $\ominus$ by $\mu \odot \nu = (\mu^{\diamond} \oplus \nu^{\diamond})^{\diamond}$ and $\mu \ominus \nu = \mu \odot \nu^{\diamond}$, then
$(MV5)$ $(\mu \oplus \nu) = (\mu^{\diamond} \odot \nu^{\diamond})^{\diamond}$,
$(MV6)$ $\mu \oplus 1 = 1$,
$(MV7)$ $(\mu \ominus \nu) \oplus \nu = (\nu \ominus \mu) \oplus \mu$,
$(MV8)$ $\mu \oplus \mu^{\diamond} = 1$,
for every $\mu, \nu \in V$. We know that $(V, \odot, 1)$ is an Abelian monoid. If we define auxiliary operations $\vee$ and $\wedge$ on V by $\mu \vee \nu = (\mu \odot \nu^{\diamond}) \oplus \nu$ and $\mu \wedge \nu = \mu \odot (\mu^{\diamond} \oplus \nu)$, for every $\mu, \nu \in V$, then $(V, \vee, \wedge, 0)$ is a lattice. This lattice is bounded and distributive. An ideal of MV-algebra V is a subset I of V, satisfying the following conditions: $(I1)$: $0 \in I$, $(I2)$: $\mu \leq \nu$ and $\nu \in I$ imply $\mu \in I$, $(I3)$: $\mu \oplus \nu \in I$, for every $\mu, \nu \in I$. A proper ideal I of V is a prime ideal of V if and only if $\mu \ominus \nu \in I$ or $\nu \ominus \mu \in I$ (or $\mu \wedge \nu \in I$ implies that $\mu \in I$ or $\nu \in I$), for every $\mu, \nu \in V$.

Definition 3. *[8] Let $A = (A, \dot{+}, ., 0, 1)$ be an MV-semiring, $V = (V, \oplus, \diamond, 0)$ be an MV-algebra, and $\phi : A \times V \longrightarrow V$ be defined by $\phi(t, v) = tv$, for every $t \in A$ and $v \in V$. Then V is called a (left) MV-semimodule over A or briefly an A-*

semimodule if for every $t, h \in A$ and $v_1, v_2 \in V$:
$(SMV1)$ if $v_1 + v_2$ is defined in V, then $tv_1 + tv_2$ is defined in V and $t(v_1 + v_2) = tv_1 + tv_2$;
$(SMV2)$ $(t \dot{+} h)v_1 = tv_1 \oplus hv_1$;
$(SMV3)$ $(t.h)v_1 = t(hv_1)$.
V is a unitary A-semimodule if A has a unity 1_A for the product, that is
$(SMV4)$ $1_A v_1 = v_1$, for every $v_1 \in V$.

III. Some results on prime A-ideals in MV-semimodules

Now, we present some results on prime A-ideals in MV-semimodules, by using some concepts as Q-ideals in semirings. Specially, we present some equivalent conditions for having prime A-ideals in MV-semimodules.

Definition 4. *[8] Consider $A = (A, \dot{+}, ., 0, 1)$ is an MV-semiring, $V = (V, \oplus, \diamond, 0)$ is an MV-algebra, V is an A-semimodule, and Ω is an ideal of V. Then Ω is called an A-ideal of V if $(I4)$: $t\gamma \in \Omega$, for every $t \in A$ and $\gamma \in \Omega$. Moreover, Ω is called a prime A-ideal of V, if $t\gamma \in \Omega$ implies that $\gamma \in \Omega$ or $t \in (\Omega : V) = \{t \in A : tV \subseteq \Omega\}$, for any $t \in A$ and $\gamma \in V$. The proper ideal P of A is called a prime ideal of A if $t.h \in P$ implies that $t \in P$ or $h \in P$, for any $t, h \in A$.*

Proposition 1. *Let A be an MV-semiring. $\rho = \{0\}$ is a prime A $-$ideal of $V \iff$*

$$Ann_A(V) = \{r \in A : rm = 0, \text{ for some } m \in V - \{0\}\}.$$

Theorem 1. *Let A be an MV-semiring. If V is an A-semimodule and Ψ is a proper A-ideal of V, then*
Ψ is a prime A-ideal of $V \iff (I\Omega \subseteq \Psi \implies I \subseteq (\Psi : V)$ or $\Omega \subseteq \Psi)$, for every ideal I of A and A-ideal Ω of V.

Proof. $(\Rightarrow)$ The proof is routine.
$(\Leftarrow)$ Let for every ideal I of A and A-ideal Ω of V, $I\Omega \subseteq \Psi$ implies that $I \subseteq (\Psi : V)$ or $\Omega \subseteq \Psi$. Let $xv \in \Psi$ and $v \notin \Psi$. Consider $I = \{z \in A : z \leq x\}$ and $\Omega = \{\sum_{i=1}^{k} t_i v : \sum_{i=1}^{k} t_i v \leq n(tv), \text{ for some } n, k \in \mathbb{N} \cup \{0\}, \text{ where } t, t_i \in A\}$. The first, we can show that I is an ideal of A. Now, we prove that Ω is an A-ideal of M:
It is clear that (I_1) is correct.
If $x \leq \sum_{i=1}^{k} t_i v \in \Omega$, for some $x \in V$, then $x = 1x \leq \sum_{i=1}^{k} t_i v \leq n(tv) \in \Omega$, for some $n \geq 0$ and $t \in A$, and so $x \in \Omega$. So (I_2) is correct.
Consider the elements of Ω: $\sum_{i=1}^{k} t_i v$ and $\sum_{i=1}^{w} s_i v$. Then there exist $n_1, n_2 \geq 0$ and $s, t \in A$ such that $\sum_{i=1}^{k} t_i v \leq n_1(tv)$ and $\sum_{i=1}^{w} s_i v \leq n_2(sv)$ and so

$$\begin{aligned} \sum_{i=1}^{k} t_i v \oplus \sum_{i=1}^{w} s_i v &\leq n_1(tv) \oplus n_2(sv) = \underbrace{tv \oplus \cdots \oplus tv}_{n_1 \text{ times}} \\ &\oplus \underbrace{sv \oplus \cdots \oplus sv}_{n_2 \text{ times}} \\ &= (t + \cdots + t)v \oplus (s + \cdots + s)v \\ &= tv \oplus sv = (t + s)v = 1(t + s)v \end{aligned}$$

It results that $\sum_{i=1}^{k} t_i v \oplus \sum_{i=1}^{w} s_i v \in \Omega$. Then (I_2) is correct. Now, for every $a \in A$ and $\sum_{i=1}^{k} t_i v \in \Omega$, there is $n \geq 0$ and $t \in A$ such that $\sum_{i=1}^{k} t_i v \leq n(tv)$. Hence

$$\begin{aligned} a(\sum_{i=1}^{k} t_i v) &\leq a(\underbrace{tv \oplus \cdots \oplus tv}_{n \text{ times}}) = a(\underbrace{tv + \cdots + tv}_{n \text{ times}}) \\ &= \underbrace{(a(tv) + \cdots + a(tv))}_{n \text{ times}} \\ &= (a.t)v + \cdots + (a.t)v \\ &= n(a.t)v. \end{aligned}$$

It means that $a(\sum_{i=1}^{k} t_i v \in \Omega$ and so (I_4) is correct.
Finally, let there is $x \in A$ and $v \in V$ such that $xv \in \Psi$ and $v \notin \Psi$.Then $y \leq x$, for every $y \in I$ and so $yv \leq xv \in \Psi$. Hence, $yv \in \Psi$ and so

$$I\Omega = \{y(\sum_{i=1}^{k} t_i v) : y, t \in A\} = \{\sum_{i=1}^{k} t_i(yv) : y, t \in A\} \subseteq \Psi.$$

Then $I \subseteq (\Psi : V)$ or $\Omega \subseteq \Psi$. Since $v \notin \Psi$, we have $I \subseteq (\Psi : V)$ and so $xV \subseteq \Psi$. Therefore, Ψ is a prime A-ideal of V. □

Lemma 1. *(i) Let A be an MV-semiring. Consider V as an A-semimodule and $\emptyset \neq \Omega \subseteq V$. Then*

$$\begin{aligned} \prec \Omega \succ = \{\varrho \in V \ : \ & \varrho \leq \varrho_1 \oplus \cdots \oplus \varrho_n \oplus a_1 r_1 \oplus \cdots \oplus a_m r_m, \\ & \text{for some} \\ & \varrho_1, \cdots, \varrho_n, r_1, \cdots, r_m \in \Omega \\ & \text{and } a_1, \cdots, a_m \in A\} \end{aligned}$$

is an A-ideal of V. It is called the A-ideal generated by Ω. In particular, for every $\alpha \in V$,

$$\prec \alpha \succ = \{x \leq n\alpha \oplus m(a\alpha) : x \in V, \ \exists n, m \in \mathbb{N} \text{ and } a \in A\}.$$

(ii) Let Ω be a proper A-ideal of V. Then $(\Omega : V) = (\Omega : \prec v \succ)$, for every $v \in V \setminus \Omega$ if and only if $\Omega = \{v \in V : Jv \subseteq \Omega\}$, for every ideal J of A such that $J \nsubseteq (\Omega : V)$.

Definition 5. *[3, 4] Let Υ be a semiring and I be an ideal of Υ. Then I is called a Q-ideal of Υ if there exists a subset Q of Υ such that:*
(1) $\Upsilon = \bigcup\{d \dot{+} I : d \in Q\}$;
(2) if $d_1, d_2 \in Q$, then $(d_1 \dot{+} I) \cap (d_2 \dot{+} I) \neq \emptyset$ if and only if $d_1 = d_2$.
Now, let $\frac{\Upsilon}{I} = \{d \dot{+} I : d \in Q\}$. Consider operations $\boxplus$ and $\boxtimes$ be defined on $\frac{\Upsilon}{I}$ as follows:

$$(d_1 \dot{+} I) \boxplus (d_2 \dot{+} I) = d_3 \dot{+} I$$

where $d_3 \in Q$ is a unique element such that $d_1 \dot{+} d_2 \dot{+} I \subseteq d_3 \dot{+} I$.

$$(d_1 \dot{+} I) \boxtimes (d_2 \dot{+} I) = d_4 \dot{+} I$$

where $d_4 \in Q$ is a unique element such that $d_1 d_2 \dot{+} I \subset d_4 \dot{+} I$. Then $(\frac{\Upsilon}{I}, \boxplus, \boxtimes, 0 \dot{+} I)$ is an MV-semiring related to I.

Note. From now on, in this paper, we let A be an MV-semiring and V be an MV-algebra.

Example 1. *Let V be a unitary A-semimodule, and Ψ be a prime A-ideal of V. Then $(\Psi : V)$ is a Q-ideal of A, if $\alpha^{\diamond}m \in \Psi$ implies $\alpha m \in \Psi$, for every $\alpha \in A$ and $m \in V$. It is routine to see that $(\Psi : V)$ is an ideal of A. By considering $Q = (A - (\Psi : V)) \cup \{0\}$, we show that $(\Psi : V)$ is a Q-ideal of A:*
It is easy to see that $A = \bigcup\{f \dot{+} (\Psi : V) : f \in Q\}$. Now, we show that $(f_1 \dot{+} (\Psi : V)) \cap (f_2 \dot{+} (\Psi : V)) = \emptyset$, for every $f_1, f_2 \in Q$, where $f_1 \neq f_2$. Let

$$(f_1 \dot{+} (\Psi : V)) \cap (f_2 \dot{+} (\Psi : V)) \neq \emptyset,$$

for some $f_1, f_2 \in Q$, where $f_1 \neq f_2$ and there exists $r \in (f_1 \dot{+} (\Psi : V)) \cap (f_2 \dot{+} (\Psi : V))$. There are 2 cases:
(I) Let $f_1 = 0$, $f_2 \neq 0$. Then there exists $t \in (\Psi : V)$ such that $r = f_2 \dot{+} t$, $rV \subseteq \Psi$, $tV \subseteq \Psi$ and $f_2V \nsubseteq \Psi$. So $rm, tm \in \Psi$, for every $m \in V$ and there is $m_1 \in V$ such that $f_2m_1 \notin \Psi$. Now, since $f_2m_1 \leq f_2m_1 \oplus tm_1 = rm_1 \in \Psi$, we have $f_2m_1 \in \Psi$ that is a contradiction. Hence $f_2 = 0$ and so $f_1 = f_2$.
(II) Let $f_1, f_2 \neq 0$, where $r = f_1 \dot{+} t_1 = f_2 \dot{+} t_2$, for $t_1, t_2 \in (\Psi : V)$. Then $f_1V \nsubseteq \Psi$, $f_2V \nsubseteq \Psi$ and there exists $m \in V$ such that $f_1m \notin \Psi$. We consider two cases:
(1) If $f_2m \in \Psi$, then

$$\begin{aligned} f_1m \leq f_1m \oplus t_1m &= (f_1 \dot{+} t_1)m = rm = (f_2 \dot{+} t_2)m \\ &= f_2m \oplus t_2m \in \Psi \end{aligned}$$

and so $f_1m \in \Psi$ that is a contradiction. So $f_1 = f_2$.
(2) Let $f_2m \notin \Psi$, too. Since $f_1 \dot{+} t_1 \leq f_2 \dot{+} t_2$, we have $(f_2 \dot{+} t_2)^{\diamond}.(f_1 \dot{+} t_1) = 0$ and so

$$f_2^{\diamond}.(f_2 \dot{+} t_2)^{\diamond}.(f_1 \dot{+} t_1) = 0.$$

It results that $f_2^{\diamond}.(f_1 \dot{+} t_1) \leq f_2^{\diamond}.t_2$ and so

$$\begin{aligned} (f_2^{\diamond}.(f_1 \dot{+} t_1))m &\leq (f_2^{\diamond}.t_2)m \\ &= f_2^{\diamond}(t_2m) \in \Psi \end{aligned}$$

Hence $(f_2^{\diamond}.(f_1 \dot{+} t_1))m \in \Psi$. Since

$$\begin{aligned} (f_2^{\diamond}.f_1)m &\leq (f_2^{\diamond}.f_1)m \oplus (f_2^{\diamond}.t_1)m = (f_2^{\diamond}.f_1 \dot{+} f_2^{\diamond}.t_1)m \\ &= (f_2^{\diamond}.(f_1 \dot{+} t_1))m \in \Psi, \end{aligned}$$

we have $f_1(f_2^{\diamond}m) = (f_2^{\diamond}.f_1)m \in \Psi$. Now, since $f_1 \notin (\Psi : V)$ and Ψ is a prime A-ideal of V, we have $f_2^{\diamond}m \in \Psi$. It follows that $f_2m \in \Psi$ that is a contradiction. Therefore, $f_1 = f_2$ and therefore, $(\Psi : V)$ is a Q-ideal of A.

Definition 6. *Consider V as an A-semimodule and Ψ as an A-ideal of V. Then Ψ is said a torsion free A-ideal, if $t\gamma = 0$ implies that $t = 0$ or $\gamma = 0$, where $t \in A$ and $\gamma \in \Psi$. In case that $V = \Psi$, V is called a torsion free MV-semimodule on A (or torsion free A-semimodule).*

Theorem 2. *Consider V as a unitary A-semimodule and Ψ as a proper A-ideal of V. If $\alpha^{\diamond}m \in \Psi$ implies $\alpha m \in \Psi$, for every $\alpha \in A$ and $m \in V$, then the following are equivalent (for comfort, set $(\Psi : V) = \Omega$) :*
(1) Ψ is a prime A-ideal of V.
(2) $\frac{V}{\Psi}$ is a torsion free $\frac{A}{\Omega}$-semimodule.
(3) $\Psi = \{m \in V : rm \in \Psi\}$, for every $r \in A \setminus \Omega$.
(4) $\Psi = \{m \in V : Jm \subseteq \Psi\}$, for every ideal J of A such that $J \nsubseteq \Omega$.
(5) $\Omega = (\Psi :\prec m \succ)$, for all $m \in V \setminus \Psi$.
(6) $\Omega = (W : L)$, for all A-ideal of V that $W \nsubseteq L$.
(7) $Ann_A(\frac{m}{\Psi}) = \Omega$, for all $m \in V \setminus \Psi$.
(8) $\Omega = \{r \in A : r\frac{m}{\Psi} = \frac{0}{\Psi}$, for some $m \in V - \{0\}\}$.

Proof. (1) $\Rightarrow$ (2) Since Ω is a Q-ideal of A, $\frac{A}{\Omega}$ is defined. Consider operation $\frac{A}{\Omega} \times \frac{V}{\Psi} \longrightarrow \frac{V}{\Psi}$ by $(a + \Omega)\frac{x}{\Psi} = \frac{ax}{\Psi}$, for every $a \in A$ and $x \in V$. Then we will have $\frac{V}{\Psi}$ is a torsion free $\frac{A}{\Omega}$-semimodule.
(2) $\Rightarrow$ (3) Consider $T = \{m \in V : rm \in \Psi\}$, for every $r \in A \setminus \Omega$. Let $m \in T$. Then $rm \in \Psi$ and so $rm = f_d(rm, 0) \in \Psi$. It results in $\frac{r}{\Omega}\frac{m}{\Psi} = \frac{rm}{\Psi} = \frac{0}{\Psi}$. Since $\frac{V}{\Psi}$ is a torsion free $\frac{A}{\Omega}$-semimodule, we have $\frac{m}{\Psi} = \frac{0}{\Psi}$ and so $m \in \Psi$. Hence $\Psi = T$.
(3) $\Rightarrow$ (4) Consider J as an ideal of A such that $J \nsubseteq \Omega$. Then there is $j \in J \setminus \Omega$. Now, let $m \in \{m \in V : Jm \subseteq \Psi\}$. Then $jm \in \Psi$ and $j \notin \Omega$ and so by (c), we see $m \in \Psi$. Hence $\Psi = \{m \in V : Jm \subseteq \Psi\}$.
(4) $\Rightarrow$ (5) By Lemma 1, the proof is clear.
(5) $\Rightarrow$ (6) Let $\Psi \nsubseteq L \subseteq V$. Then there is $m \in L \setminus \Psi$ and so by (5), we have $\Omega = (\Psi :\prec m \succ)$. Since $m \in L$ and $\Omega = (\Psi : V)$, we see $(\Psi : L) = \Omega$.
(6) $\Rightarrow$ (7) Consider $m \in V \setminus \Psi$. Let $r \in Ann_A(\frac{m}{\Psi})$. Then $r\frac{m}{\Psi} = \frac{0}{\Psi}$ and so $rm \in \Psi$. We consider $L =\prec m \succ$ and so by (6), $(\Psi : L) = \Omega$. We show that $rL \subseteq \Psi$. Let $x \in L$. Then $x \leq tm \oplus k(\alpha m)$, for some integers $t, k \geq 0$ and $\alpha \in A$ and so

$$rx \leq r(tm) \oplus rk(\alpha m) = t(rm) \oplus k\alpha(rm) \in \Psi.$$

It results in $rx \in \Psi$ and so $r \in (L : \Psi) = \Omega$. Hence $Ann_A(\frac{m}{\Psi}) \subseteq \Omega$. We can see that $\Omega \subseteq Ann_A(\frac{m}{\Psi})$ and so $Ann_A(\frac{m}{\Psi}) = \Omega$.
(7) $\Rightarrow$ (8) We have

$$\begin{aligned} \{r \in A : r\frac{m}{\Psi} &= \frac{0}{\Psi}, \text{ for some } \frac{m}{\Psi} \in \frac{V}{\Psi} \setminus \{\frac{0}{\Psi}\}\} \\ &= \{r \in A : f_d(rm, 0) \in \Psi, \text{ for some } m \in V \setminus \Psi\} \\ &= \{r \in A : rm \in \Psi, \text{ for some } m \in V \setminus \Psi\} \\ &= Ann_A(\frac{m}{\Psi}) \\ &= \Omega \end{aligned}$$

(8) $\Rightarrow$ (1) It is easy. □

IV. Conclusion

In [8], the new definition of MV-modules was introduced . After that, we used that definition to work on modular structures. In this paper, some equivalent conditions for having prime A-ideals in MV-semimodules were presented. A number of theorems were proved that could be proved by the old definitions only under certain conditions. We intend to expand the use of this definition.

REFERENCES

[1] R. Cignoli, I. M. L. $D^{'}$Ottaviano and D. Mundici, " Algebraic foundations of many-valued reasoning", Kluwer Academic, Dordrecht, 2000.

[2] A. Di Nola and C. Russo, " Semiring and semimodules issues in MV-algebras", Communications in Algebra, vol. 41, pp. 1017–1048, 2013.

[3] V. Gupta and N. Chaudhari, "Some results on semirings", Radovi Mathematiki, vol. 12, pp. 13–18, 2003.

[4] V. Gupta and N. Chaudhari, "Right π-regular semirings", Sarajevo Journal of Mathematics, vol. 14, pp. 3–9, 2006.

[5] S. Saidi Goraghani and R. A. Borzooei, "On injective MV-modules", Bulletin of of the Section of Logic, vol. 47, pp. 283–298, 2018.

[6] S. Saidi Goraghani and R. A. Borzooei, "Results on prime ideals in PMV-algebras and MV-modules", Italian Journal of Pure and Applied Mathematics, vol. 37, pp. 183–196, 2017.

[7] S. Saidi Goraghani and R. A. Borzooei, "Most results on A-ideals in MV-moduls", Journal of Algebraic Systems, vol. 5, pp. 1–13, 2017.

[8] S. Saidi Goraghani, R. A. Borzooei and S. S. Ahn, "New definition of MV-semimodules", Journal of Intelligent and Fuzzy System, vol. 41, pp. 3715–3726, 2021.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Notes on energy of matching in fuzzy graphs

1st Majid Khalili, 2nd Rajab Ali Borzooei

1st Department of Mathematics, Chabahar Branch, Islamic Azad University, Chabahar, Iran, majidkhalili@iauchab.ac.ir
2nd Department of Mathematics, Shahid Beheshti University, Tehran, Iran, borzooei@sbu.ac.ir

***Abstract*— In this paper, by defining algebraic structures on a matching in fuzzy graphs, we will introduce the energy of matching. These structures will begin by defining specific incidence matrices that are built on the matching. In the algebraic properties of these definitions, relationships with fuzzy numbers of a matching will be observed.**

***Index Terms*— energy, incidence matrix, matching edge number, matching vertex number, generator sequence**

I. Introduction

Euler's solution of the Konigsberg bridges' problem in 1753, can be considered the beginning of graph theory [17]. The results he obtained in the study of this issue established the main features of the graph theory. Graphs are visually a set of points connected by arcs. Solving graph problems required that they be stored in computers memory. For this purpose, many matrices are introduced based on graphs. A concept on one of these matrices, the adjacency matrix, is called the spectrum of graph [9]. Due to the practical aspects of this subject, extensive studies have been conducted on them in physics, chemistry, computer science and other branches of mathematics. The concept that relates to the spectrum of a graph is energy. Many mathematical properties are studied in the energy of graphs. Energy values are discussed in [5] and [25]. Bounds on energy [8], [13], [21], hyperenergetic, hypoenergetic and equienergetic graphs [4], [10], [15], [27], and energy of different graphs [12], [14], [16], [30], is under study.

The fuzzy set was defined by Zadeh in 1965 [31]. His goal was to develop a theory for ambiguous sets, which is characteristic of most real-world sets. Like crisp sets, relationships between fuzzy subsets are defined by fuzzy relations. Just as graphs are representations of relationships, fuzzy relationships are also represented by fuzzy graphs. The fuzzy graph was first introduced by Rosenfeld [29] and has been followed by applied studies in this concept [3], [7], [22], [28]. Narayanan and Mathew [23] defined the concept of energy for fuzzy graphs based on the same concept in weighted graphs. Energy is also defined for other specific fuzzy graphs [2], [11], [18], [24]. One of the topics studied in graph theory and computer science is matching. Sumosandaram first defined matching in fuzzy graphs based on the concept of domination. Today, domination is discussed independently in fuzzy graphs [1]. Matching was then introduced in a fuzzy graph using the concept of fractional matching [26]. The latest definition of matching in fuzzy graphs, while retaining its properties in crisp graphs, has varied matching based on edge and vertex values [19]. In this paper, based on the recent view of matching in fuzzy graphs, we have introduced energy of matching as another tool for comparing matchings.

II. Preliminaries

Definition 1. *[29]. If $G^* = (V, E)$ is a crisp graph, $G = (V, \sigma, \mu)$ is called a fuzzy graph on G^* when $\sigma : V \to [0, 1]$, $\mu : E \to [0, 1]$ and for all u and v in V, $\mu(uv) \leq \sigma(u) \wedge \sigma(v)$.*

Definition 2. *[20]. Let $G = (V, \sigma, \mu)$ be a fuzzy graph. A fuzzy subgraph of G such as $M = (W, \sigma_M, \mu_M)$, is a matching in G, if for all $u \in W$, only one $v \in W$ can be found that $u \neq v$ and $\mu_M(uv) > 0$.*

Definition 3. *[20]. Let G be a fuzzy graph on G^* and M be a matching on G. Then,*

(i) the matching edge number of M, is defined by,

$$\lambda_E(M) = \sum_{e \in E_s(M)} \mu(e),$$

(ii) the matching vertex number of M, is defined by,

$$\lambda_V(M) = \sum_{v \in V_s(M)} \sigma(v),$$

(iii) the matching crisp number of M is define by $\lambda_C(M) = |E_s(M)|$.

As we know, if the eigenvalues of a graph G are $l_0 > l_1 > ... > l_{t-1}$, and their multiplicities are $m(l_0), m(l_1), ..., m(l_{t-1})$, then the spectrum of G is

$$Spec(G) = \begin{pmatrix} l_0 & l_1 & ... & l_{t-1} \\ m(l_0) & m(l_1) & ... & m(l_{t-1}) \end{pmatrix}.$$

The energy of a graph Γ, denoted by $E(\Gamma)$, is defined by

$$E(\Gamma) = \sum_{i=1}^{n} |l_i|.$$

Here n is the number of vertices in Γ, and $l_i, 1 \leq i \leq n$ are the eigenvalues of Γ [6], [13].

III. Edge matching occurrence matrix

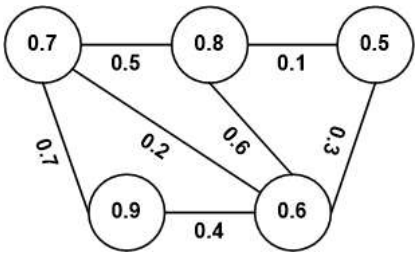
Let $G = (V, \mu, \sigma)$ be a fuzzy graph and M a matching in G such that $M = (V_S(M), \mu, \sigma)$. If $|V_S(M)| = m$, by re-labeling and based on pairwise separate edges, a finite sequence of regular triplets can be made as follows,

$$v_1e_1v_2, v_3e_2v_4, v_5e_3v_6, ..., v_{m-1}e_{\frac{m}{2}}v_m.$$

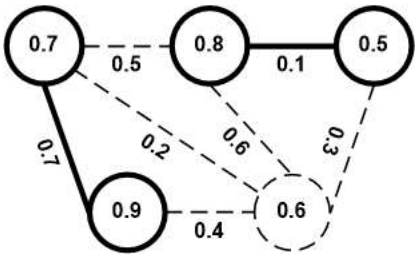
This sequence is called *the matching generator sequence*. Therefore, we will have the following display for a matching, based on its generator sequence,

$$M = \langle v_{2k-1}e_kv_{2k}\rangle \quad , \quad 1 \leq k \leq \frac{m}{2}$$

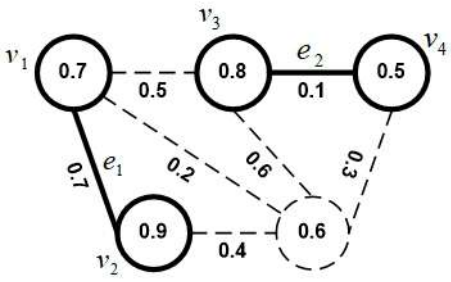
Example 1. *In the Figure 1, we see a fuzzy graph, a matching of it, and a re-labeling to introduce the matching generator sequence.*



(a) $G_1 = (V, \sigma, \mu)$.



(b) A matching M_1 in G_1.



(c) A labeling for $M_1 = \langle v_1e_1v_2, v_3e_2v_4\rangle$.

Fig. 1: Description of a matching with the generator sequence.

Definition 4. *Let $G = (V, \sigma, \mu)$ be a fuzzy graph, M a matching in G such that $M = (V_S(M), \mu, \sigma)$, and*

$$M = \langle v_1e_1v_2, v_3e_2v_4, v_5e_3v_6, ..., v_{m-1}e_{\frac{m}{2}}v_m\rangle.$$

The edge matching occurrence matrix of G is a matrix like $A_{\mu(M)} = [\mu_{ij}]_{m\times m}$ in which,

$$\mu_{ij} = \begin{cases} \mu_{i+1j+1} = \mu(e_{\frac{i+1}{2}}) & i = j = 2k+1, 0 \leqslant k \leqslant \frac{m}{2} - 1 \\ 0 & i \neq j \end{cases}$$

Example 2. *Consider matching M_1 in Example 1. The edge matching occurrence matrix for M_1 is*

$$A_{\mu(M_1)} = \begin{bmatrix} 0.7 & 0 & 0 & 0 \\ 0 & 0.7 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 0.1 \end{bmatrix}.$$

Let $M = \langle v_1e_1v_2, v_3e_2v_4, v_5e_3v_6, ..., v_{m-1}e_{\frac{m}{2}}v_m\rangle$ be a matching in G. Then *the matching defeated generator sequence* of M is defined as follow,

$$v_2e_1v_1, v_4e_2v_3, v_6e_3v_5, ..., v_me_{\frac{m}{2}}v_{m-1}.$$

Therefore, based on the defeated generator sequence, the description of M is as follows,

$$M = \langle v_{2k}e_kv_{2k-1}\rangle \quad , \quad 1 \leq k \leq \frac{m}{2}.$$

As a result, the edge matching occurrence matrix can be described as a matrix that its columns are labeled by the vertices in generator sequence (in the same order), and its rows are labeled by the vertices in defeated generator sequence (in the same order). Each matrix element is the fuzzy value of edge connects the vertices in this labeling. With this explanation, matrix $A_{\mu(M_1)}$ rows and columns are labeled as follows,

$$A_{\mu(M_1)} = \begin{array}{c} \\ v_2 \\ v_1 \\ v_4 \\ v_3 \end{array} \begin{array}{c} \begin{array}{cccc} v_1 & v_2 & v_3 & v_4 \end{array} \\ \begin{bmatrix} 0.7 & 0 & 0 & 0 \\ 0 & 0.7 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 0.1 \end{bmatrix} \end{array}$$

Proposition 1. *Let $G = (V, \sigma, \mu)$ be a fuzzy graph, M a matching in G such that $M = (V_S(M), \mu, \sigma)$, and $A_{\mu(M)} = [\mu_{ij}]_{m\times m}$ be edge fuzzy matching occurrence matrix of M. Then,*

$$trace(A_{\mu(M)}) = 2\lambda_E(M).$$

Proof: Consider

$$M = \langle v_1e_1v_2, v_3e_2v_4, v_5e_3v_6, ..., v_{m-1}e_{\frac{m}{2}}v_m\rangle,$$

then

$$trace(A_{\mu(M)}) = \sum_{i=1}^{m} \mu_{ii} = \sum_{i=0}^{\frac{m}{2}-1} \mu_{2i+12i+1} + \sum_{j=1}^{\frac{m}{2}} \mu_{2j2j}.$$

According to the Definition 4,

$$\mu_{2j2j} = \mu(e_j) \quad , \quad 1 \leq j \leq \frac{m}{2}$$

$$\mu_{2i+12i+1} = \mu\left(e_{\frac{(2i+1)+1}{2}}\right) = \mu(e_{i+1}) \ , \ 0 \leq i \leq \frac{m}{2} - 1,$$

but,

$$\mu(e_j)|_{1\leq j\leq \frac{m}{2}} = \mu(e_{i+1})|_{0\leq i\leq \frac{m}{2}-1},$$

so,

$$\begin{aligned} trace(A_{\mu(M)}) &= \sum_{i=0}^{\frac{m}{2}-1} \mu(e_{i+1}) + \sum_{j=1}^{\frac{m}{2}} \mu(e_j) \\ &= \sum_{j=1}^{\frac{m}{2}} \mu(e_j) + \sum_{j=1}^{\frac{m}{2}} \mu(e_j) \\ &= \lambda_E(M) + \lambda_E(M) = 2\lambda_E(M). \end{aligned}$$

■

Proposition 2. *Let $G = (V, \sigma, \mu)$ be a fuzzy graph, M a matching in G, and $A_{\mu(M)} = [\mu_{ij}]_{m\times m}$ be edge matching occurrence matrix of M.*

(i) *Eigenvalues of the matrix $A_{\mu(M)}$ are exactly the edge values of the matching M.*

(ii) *The maximum number of eigenvalues of $A_{\mu(M)}$ is $\frac{m}{2}$, and multiplicity of each eigenvalue is at least 2.*

(iii) *If the edge values of M are distinct, then $A_{\mu(M)}$ exactly has $\frac{m}{2}$ eigenvalues and multiplicity of each eigenvalue is 2.*

Proof: Clearly, all cases result from the diagonal property of matrix $A_{\mu(M)}$. Note that in (ii), because the edge values may not be distinct, the number of eigenvalues decreases and the multiplicity increases. ■

Corollary 1. *Let $G = (V, \sigma, \mu)$ be a fuzzy graph, M be a matching in G, and $A_{\mu(M)} = [\mu_{ij}]_{m\times m}$ be edge matching occurrence matrix of M. If the distinctive edge values of the matching M are $\mu_1 > \mu_2 > \mu_3 > ... > \mu_l$, then*

(i)

$$Spec(M) = \begin{pmatrix} \mu_1 & \mu_2 & ... & \mu_l \\ m_1 & m_2 & ... & m_l \end{pmatrix}$$

such that $m_i \geq 2$ for all $1 \leq i \leq l$.

(ii) *If all of the edge values of M are distinct, then*

$$Spec(M) = \begin{pmatrix} \mu_1 & \mu_2 & ... & \mu_{\frac{m}{2}} \\ 2 & 2 & ... & 2 \end{pmatrix}$$

IV. Vertex matching occurrence matrix

Definition 5. *Let $G = (V, \sigma, \mu)$ be a fuzzy graph, and M a matching in G such that, $M = \langle v_1e_1v_2, v_3e_2v_4, v_5e_3v_6, ..., v_{m-1}e_{\frac{m}{2}}v_m\rangle$. The vertex matching occurrence matrix of G is a matrix*

$$A_{\sigma(M)} = [\sigma_{ij}]_{m\times m}$$

in which, for $i = j = 2k + 1(0 \leqslant k \leqslant \frac{m}{2} - 1)$,

$$\sigma_{ij} = \sigma_{i+1j+1} = \sigma(v_i) + \sigma(v_{i+1})$$

and for $i \neq j$, $\sigma_{ij} = 0$.

Example 3. *With the assumptions of Example 1, the vertex matching occurrence matrix of M_1 is as follows,*

$$A_{\sigma(M_1)} = \begin{array}{c} \\ v_2 \\ v_1 \\ v_4 \\ v_3 \end{array} \begin{array}{c} \begin{array}{cccc} v_1 & v_2 & v_3 & v_4 \end{array} \\ \begin{bmatrix} 1.6 & 0 & 0 & 0 \\ 0 & 1.6 & 0 & 0 \\ 0 & 0 & 1.3 & 0 \\ 0 & 0 & 0 & 1.3 \end{bmatrix} \end{array}$$

As it is seen, the vertex fuzzy matching occurrence matrix can be described as a matrix that its columns are labeled by the vertices in generator sequence (in the same order), and its rows are labeled by the vertices in defeated generator sequence (in the same order). Each matrix element is the sum of the fuzzy values of the vertices that are connected together.

Proposition 3. *Let $G = (V, \sigma, \mu)$ be a fuzzy graph, M be a matching in G, and $A_{\sigma(M)} = [\sigma_{ij}]_{m\times m}$ be the vertex matching occurrence matrix of M. Then,*

$$trace(A_{\sigma(M)}) = 2\lambda_V(M).$$

Proof: The argument is similar to the Proposition 1. ■

Proposition 4. *Let $G = (V, \sigma, \mu)$ be a fuzzy graph, M be a matching in G, and $A_{\sigma(M)} = [\sigma_{ij}]_{m\times m}$ be the vertex matching occurrence matrix of M.*

(i) *Eigenvalues for matrix $A_{\sigma(M)}$ are*

$$\lambda_i = \lambda_{i+1} = \sigma(v_i)+\sigma(v_{i+1}), i = 2k+1, 0 \leq k \leq \frac{m}{2}-1$$

(ii) *The maximum number of eigenvalues of $A_{\sigma(M)}$ is $\frac{m}{2}$, and multiplicity of each eigenvalue is at least 2.*

Proof: Similar to Proposition 2 all cases result from the diagonal property of matrix $A_{\sigma(M)}$. ■

Considering the properties of the vertex and edge adjacency matrices of an matching, we define two types of energy as follows.

Definition 6. *Let $G = (V, \sigma, \mu)$ be a fuzzy graph, $M = (V_S(M), \sigma_M, \mu_M)$ a matching, $A_{\mu(M)}$ the edge matching occurrence matrix and $A_{\sigma(M)}$ be the vertex matching occurrence matrix of M,*

(i) *The edge energy of M is defined as follows,*

$$E_E(M) = 2\lambda_E(M)$$

(ii) *The vertex energy of M is defined as follows,*

$$E_V(M) = 2\lambda_V(M).$$

Because according to the above definition, the matching energy is related to edge and vertex numbers, comparisons between matchings and the sizes of their components can be read based on the matching energy. In the first step it can be concluded that the vertex energy of a matching is always greater than its edge energy, in other words

$$E_V(M) > E_E(M).$$

In general, in all matching optimization situations, we have reached a matching with maximum energy or minimum energy.

V. Conclusions

Generator and defeated generator sequences can be other tools for expressing fuzzy matchings. These sequences play a key role in representing matching occurrence matrices diagonally. Matching numbers are directly related to the eigenvalues of these matrices. Therefore, the energy of the matchings can be expressed in terms of fuzzy numbers, and the comparison of the optimal states of the matchings can be expressed in terms of energy.

References

[1] S. Afsharmanesh and R. A. Borzooei, "Domination in fuzzy incidence graphs based on valid edges, " Journal of Applied Mathematics and Computing, vol. 67, 2021, DOI:10.1007/s12190-021-01510-3.

[2] M. Akram, D. Saleem, and B. Davvaz, "Energy of double dominating bipolar fuzzy graphs, " J. Appl. Math. Comput, vol. 61, pp. 219-234, 2019.

[3] M. Akram, M. Sarwar and R. A. Borzooei, "A novel decision-making approach based on hypergraphs in intuitionistic fuzzy environment,"Journal of Intelligent and Fuzzy Systems, vol. 35, pp. 1905-1922, August 2018.

[4] R. Balakrishnan, "The energy of a graph, " Linear Algebra Appl. vol. 387, pp. 287-295, 2004.

[5] R. B. Bapat and S. Pati, "Energy of a graph is never an odd integer," Bull. Kerala Math. Assoc. vol. 1, pp. 129-132, 2004.

[6] N. Biggs, "Algebraic graph theory," Cambridge University Press, 1974.

[7] R. A. Borzooei and H. Rashmanlou, "New concepts of vague graphs, "International Journal of Machine Learning and Cybernetic, vol. 8, pp. 1081-1092, August 2017.

[8] R. Brualdi, "Energy of a graph, Notes to AIM Workshop on spectra of families of matrices described by graphs, digraphs and sign patterns," 2006.

[9] L. Collatz, and U. Sinogowitz, "Spektren endlicher Grafen." Abh. Math. Sem. Univ. Hamburg vol. 21, pp. 63-77, 1957.

[10] D. M. Cvetkovic and I. Gutman, "Applications of Graph spectra," Math. Inst. Belgrade, 2009.

[11] G. Deepa, B. Praba and V. M. Chandrasekaran, "A Study on Energy of an Intuitionistic Fuzzy Directed Graph, " Research J. Pharm. and Tech, vol. 9, pp. 190-195, 2016.

[12] W. Du, X. Li and Y. Li, "Various energies of random graphs," MATCH Commun. Math. Comput.Chem, vol. 64, pp. 251-261, 2010.

[13] I. Gutman, "The energy of a graph: old and new results, in: A. Betten, A. Kohnert, R. Laue, A. Wassermann (Eds.)," Proceedings on the conference on Algebraic Combinatorics and Applications, Berlin, Springer-Verlag, pp. 196-211 2001.

[14] I. Gutman and S. Zare Firoozabadi, J. A. de la Pena, J. Rada, On the energy of regular graphs, MATCH Commun. Math. Comput. Chem, vol. 57, pp. 435-442, 2007.

[15] G. Indulal and A. Vijayakumar, "On a pair of equienergetic graphs," MATCH Commun. Math. Comput. Chem. vol. 55, pp. 83-90, 2006.

[16] G. Indulal and A. Vijayakumar,"Energies of some non-regular graphs," J. Math. Chem, vol. 42, pp. 377-386, 2007.

[17] R. J. Wilson, "History of Graph Theory," in: Jonathan L. Gross, Jay Yellen (Eds.), Handbook of Graph theory, CRC Press, 2004.

[18] A. Kalimulla, R. Vijayaragavan and S. Sharief Basha, "Dominating Energy of Operations on Intuitionistic Fuzzy Graphs, " International Journal of Engineering & Technology, vol. 7, pp. 328-335, 2018.

[19] M. Khalili, R. A. Borzooei and M. Deldar, "Matching numbers in fuzzy graphs, " Journal of Applied Mathematics and Computing, vol. 67, pp. 1-22, 2021.

[20] M. khalili and R. A. Borzooei, "New concepts of matching in fuzzy graphs," 7th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS), 2019 Jan 29-31, Bojnord, Iran, IEEE, 2019.

[21] H. Liu, M. Lu and F. Tian, "Some upper bounds for the energy of graphs," J. Math. Chem. vol. 41, pp. 45-57, 2007.

[22] S. Mathew, J. N. Mordeson and D. S. Malik, "Fuzzy graph theory, "Springer International Publishing, 2018.

[23] A. Narayanan and S. Mathew, "Energy of a fuzzy graph, " Annals of fuzzy mathematics and Informatics, vol. 6, pp. 455-465, 2013.

[24] A. Nagarani, and S. Vimala, "Energy of Fuzzy Regular and Graceful Graphs, " Asian Research Journal of Mathematics, vol. 4, pp. 1-8, 2017.

[25] S. Pirzada and I. Gutman, "Energy of a graph is never the square root of an odd integer," Appl.Anal. Discrete Math. vol. 2, pp. 118-121, 2008.

[26] P. V. Ramakrishnan and M. Vaidyanathan, "Matching in fuzzy graphs, " Proceedings of National Conference on Discrete Mathematics and its Applications, NCDMA 2007, Madurai, Tamil Nadu, India, September, pp. 27–29, 2007.

[27] H. S. Ramane, H. B. Walikar, S. B. Rao, B. D. Acharya, P. R. Hampiholi, S. R. Jog and I. Gutman, "Equienergetic graphs," Kragujevac J. Math. vol. 26, pp. 5-13, 2004.

[28] H. Rashmanlou and R. A. Borzooei, "New concepts of interval-valued intuitionistic (S,T)-fuzzy graphs," Journal of Intelligent and Fuzzy Systems, vol. 30, pp. 1893-1901, November 2016.

[29] A. Rosenfeld, "Fuzzy graphs," Fuzzy Sets and Their Applications to Cognitive and Decision Processes, Academic Press, New York, pp. 77-95, 1975.

[30] I. Shparlinski, "On the energy of some circulant graphs," Linear Algebra Appl, vol. 414, pp. 378-382, 2006.

[31] L.A. Zadeh, "Fuzzy sets," Information and Control, vol. 8, pp. 338-353, 1965.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Numerical solution for Interval Initial Value Problems based on interactive arithmetic

1st Estevão Esmi, 2nd Cristina Sacilotto, 3rd Vinícius F. Wasques and 4th Laécio C. Barros

1st, dept. Applied Mathematics, *University of Campinas*, Campinas, Brazil, eelaureano@gmail.com
2nd dept. Applied Mathematics, *University of Campinas*, Campinas, Brazil, cristinasacilotto@gmail.com
3rd dept. Mathematics, *São Paulo State University*, Rio Claro, Brazil, vwasques@outlook.com
3rd Ilum School of Science, *Brazilian Center for Research in Energy and Materials*, Campinas, Brazil, vwasques@outlook.com
4th dept. Applied Mathematics, *University of Campinas*, Campinas, Brazil, laeciocb@ime.unicamp.br

***Abstract*— This work presents a study of Interval Initial Value Problems (IIVPs), where the derivative is given by the generalized Hukuhara derivative (gH-derivative) and the initial condition is given by an interval. The focus of the paper is to provide the numerical approximations for the solutions associated with the gH-derivative of IIVPs. This article considers the Euler numerical method, where the classical arithmetic operation presented in the method is adapted for intervals. The arithmetic considered here is obtained from the sup-J extension principle, where J is a particular family of joint possibility distributions. This family gives raise to different types of interactivity and this work shows what kind of interactivity is necessary in the numerical method, in order to approximate the solution via gH-derivative. To illustrate the results, the paper focuses in the decay Malthusian model.**

***Index Terms*— Interval initial value problem, Generalized Hukuhara derivative, Euler method, Malthusian model, Interactivity**

I. Introduction

Ordinary and Partial Differential equations (ODEs and PDEs) are fields of Mathematics which are widely studied mainly by mathematicians, physicists and engineers. This area can be used to described several phenomenon of nature, such as physical, chemical and biological processes [1].

The differential equation models the behaviour of the processes while the additional conditions, such as the initial quantity of individuals in a population study, are given by initial or boundary conditions. These problems are called Initial Value Problems (IVPs) or Boundary Value Problems (BVPs).

Although nowadays there are several mathematical methods for solving differential equations, the task of solving ODEs and PDEs are not so easy. Hence, numerical methods may be useful tools to study such problems. In order these numerical methods be consistent, the outputs obtained from them must approximate the analytical solution to the problem, that is, the analytical and numerical solutions must be coherent.

A simple and well-known numerical method is Euler's method, which approximates the analytical solution to the IVP. This method is used not only to study classical differential equations, but also considered to study Fuzzy Differential Equations (FDEs). A FDE can be classified as a differential equation that presents some parameter or state with fuzzy values. For example, the derivatives in the differential equation can be given by a fuzzy derivative (such as the Hukuhara derivative and its generalizations [2, 3] or interactive derivatives [4, 7]) or the parameters and additional conditions can be given by fuzzy numbers [5, 6].

In the context of fuzzy set theory the numerical methods, such as Euler and Runge-Kutta methods, are being considered to provide numerical approximations of analytical solutions to IVPs based on some metric (usually Hausdorff metric [8]). In general this study is given by applying the classical numerical method to the endpoints of fuzzy numbers [9]. Other approach is extending the classical arithmetic operations in the method for fuzzy numbers by the Zadeh and sup-J extension principles [10].

A particular case of FDE is an Interval Initial Value Problem (IIVP), where the variables present in the problem are given by intervals instead of fuzzy numbers. The same tools of fuzzy set theory can be employed for interval theory, since intervals are particular cases of fuzzy numbers.

This paper focuses on the study of IIVPs, where the derivative is given by the generalized Hukuhara derivative (gH-derivative) and the initial condition is given by an interval. In this study the numerical Euler's method is considered in order to compare the analytical solution obtained from the two types of the gH-derivative with the numerical solution based on interactive arithmetic, whose interactivity is associated with a parametrized family of Joint Possibility Distributions (JPDs), denoted by J_γ [11].

This work shows which types of interactivity are necessary to obtain a consistent numerical solution with the analytical solution to the problem. In order to illustrate the results, the

decay Malthusian model is considered.

The article is divided as follows. Section II provides the background for this study, bringing basics concepts of the classical and fuzzy/interval set theories. Section III presents the numerical approximations for IIVP, which is the main result of the paper. Section IV presents an example for a decay Malthusian model. Finally, Section V provides the final remarks of the article.

II. Mathematical Background

This section brings the mathematical concepts required for this paper.

A. Classical set theory

We start this section introducing the concepts of Initial Value Problem (IVP) which consists in an ordinary differential equation and initial condition given by

$$\begin{cases} \frac{dy_i}{dt} = f(t, y_1, y_2, \ldots, y_n) \\ y(t_0) = y_0 \in \mathbb{R}^n \end{cases} \tag{1}$$

where f_i is a function that depends on $y_1, y_2, ..., y_n$ and t and each $y_i : \mathbb{R} \to \mathbb{R}^n$, for $i = 1, \ldots, n$ denotes functions that depend on time t.

It is possible to obtain a solution for IVPs through the theory of differential equations which can be a complicated task depending on the equation that describes the field.

Numerical methods can be an alternative approch to study an IVP. For example Euler's method approximates the solution of an IVP since it consists in determining numerical solutions for (II-A). The algorithm of the method is given by

$$y_i^{(k+1)} = y_i^{(k)} + hf_i(t_k, y_1^{(k)}, y_2^{(k)}, \ldots, y_n^{(k)}) \tag{2}$$

with $0 \leq k \leq N-1$, where N is the number of partitions of the interval time divided in equally spaced intervals $[t_k, t_{k+1}]$ with size h and initial condition $(t^0, y_1^{(0)}, y_2^{(0)}, \ldots, y_n^{(0)})$.

For the unidimensional case, Euler's method given by

$$y^{(k+1)} = y^{(k)} + hf(t_k, y(k)), \tag{3}$$

approximates the initial value problem

$$\begin{cases} \frac{dy}{dt} = f(t, y) \\ y(t_0) = y_0 \in \mathbb{R} \end{cases},$$

whit local error $\mathcal{O}(h^2)$.

B. Fuzzy set theory

This subsection presents the fuzzy set theory and a mathematical background necessary for this paper. The fuzzy set theory was introduced by Zadeh in 1965 [12]. A fuzzy set and its α-levels are defined as follows.

Definition 1. *A subset A of a universe X is said to be a fuzzy subset if it is characterized by the function $\varphi_A : X \to [0, 1]$, which is called by the membership function of A.*

Definition 2. *The α-levels of A are defined as the classical sets*

$$[A]^\alpha = \{x \in X : \varphi_A(x) \geq \alpha\}$$

for $0 < \alpha \leq 1$ and for $\alpha = 0$ it is defined by

$$[A]^0 = cl\{x \in X : \varphi_A(x) > 0\},$$

if X is a topological space and the symbol cl stands for the closure of a set.

A fuzzy number is defined by the following definition [11].

Definition 3. *Let A be a fuzzy subset of $\mathbb{R}$ with membership function $\varphi_A : \mathbb{R} \to [0, 1]$. So A is a fuzzy number if it satisfies the following properties:*

- *All of α-levels of A are not empty, with $0 \leq \alpha \leq 1$;*
- *All of α-levels of A are closed intervals of $\mathbb{R}$;*
- *$SuppA = \{x \in \mathbb{R} : \varphi_A(x) > 0\}$ is bounded.*

Since the α-levels of a fuzzy number A are intervals, we represent them by $[A]^\alpha = [a_\alpha^-, a_\alpha^+]$. A translation of a fuzzy number A by the midpoint of $[A]^1$, that is $a = 0.5(a_1^- + a_1^+)$, is defined by $[A^t]^\alpha = [a_\alpha^{-t}, a_\alpha^{+t}]$. From this translation, we emphasize two interesting properties: $a_\alpha^{-t} \leq 0 \leq a_\alpha^{+t}$, for all $\alpha \in [0, 1]$ and $[A]^\alpha = [A^t]^\alpha + a$. This concept can be also applied to intervals, as the next example illustrate.

Example: Let $A = [\underline{a}, \overline{a}] = [2, 4]$ be an interval. The translated version of A by its midpoint $a = 3$ is defined by the interval $A^t = [\underline{a}^t, \overline{a}^t] = [-1, 1]$. Note that $A = A^t + a$ and $\underline{a}^t \leq 0 \leq \overline{a}^t$.

Next we provide the definition of sup-J extension principle, which is used here to extend arithmetic operations for fuzzy numbers, in particular for intervals.

Definition 4. *Let $J \in \mathcal{F}(\mathbb{R}^n)$ be a joint possibility distribution of $(A_1, \ldots, A_n) \in \mathbb{R}_\mathcal{F}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$. The sup-$J$ extension of f at $(A_1, \ldots, A_n)$, is the fuzzy set whose membership function is given as follows:*

$$f_J(A_1, \ldots, A_n)(y) = \bigvee_{(x_1, \ldots, x_n) \in f^{-1}(y)} J(x_1, \ldots, x_n), \tag{4}$$

where the symbol $\bigvee$ stands for the supremum operator and $f^{-1}(y) = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : f(x_1, \ldots, x_n) = y\}$.

The sup-J extension principle is in fact a generalization of Zadeh's extension principle for multiple variables, since these extensions are equal if $J(x_1, \ldots, x_n) = J_\wedge(x_1, \ldots, x_n) = \varphi_{A_1}(x_1) \wedge \ldots \wedge \varphi_{A_n}(x_n)$, where the symbol $\wedge$ represents the minimum operator.

Next, the definition of interactive fuzzy variables is presented.

Definition 5. *Let $A_1, \ldots, A_n \in \mathbb{R}_\mathcal{F}$ and J be their joint possibility distribution. The fuzzy numbers $A_1, \ldots, A_n$ are called J-interactive, or simply interactive, if $J \neq J_\wedge$.*

The generalized Hukuhara difference (gH-difference) and the associated derivative are presented as follows.

Definition 6. *Given two fuzzy numbers $A, B \in \mathbb{R}_\mathcal{F}$ the generalized Hukuhara difference (gH-difference for short) is the fuzzy number C, if it exists, such that*

$$A -_{gH} B = C \iff \begin{cases} (I)\ A = B + C \text{ or} \\ (II)\ B = A - C \end{cases}, \tag{5}$$

where the symbols "+" and "-" represent the fuzzy standard sum and difference, respectively.

Definition 7. *The function $f : (a, b) \to \mathbb{R}_{\mathcal{F}}$ is generalized Hukuhara differentiable (gH-differentiable) at t_0 if the limit*

$$\lim_{h \to 0} \frac{f(t_0 + h) -_{gH} f(t_0)}{h}$$

exists and it is equal to some element $f'_{gH}(t_0) \in \mathbb{R}_{\mathcal{F}}$ at metric d_∞. The number $f'_{gH}(t_0)$ is called by the gH-derivative of f at t_0.

If the gH-difference of type (I) is considered in the above limit, then the function f is said to be gH-differentiable of type (I), or simply (I)-gH-differentiable. On the other hand, if the gH-difference of type (II) is considered in the above limit, then the function f is said to be gH-differentiable of type (II), or simply (II)-gH-differentiable.

This paper focuses on a parametrized family of joint possibility distributions $\{J_\gamma \mid \gamma \in [0, 1]\}$ that was employed by Esmi et al. [11]. Although the construction is very technical, these distributions can be applied for any pair of fuzzy numbers $A_1, A_2 \in \mathbb{R}_{\mathcal{F}_c}$, in particular for any pair of intervals. The construction of these distributions will be omitted here, but for illustration, one can see in Figures 1, 2 and 3 the representations of the joint possibility distributions J_γ between intervals $A = [\underline{a}, \overline{a}]$ and $B = [\underline{b}, \overline{b}]$ for $\gamma = 1$, $\gamma = 0.5$ and $\gamma = 0$, respectively.

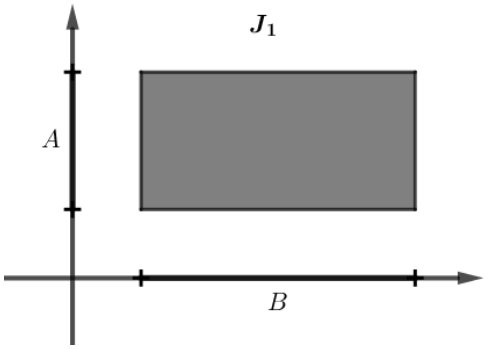


Fig. 1. Graphical representation of the joint possibility distributions J_1 of the intervals A and B.

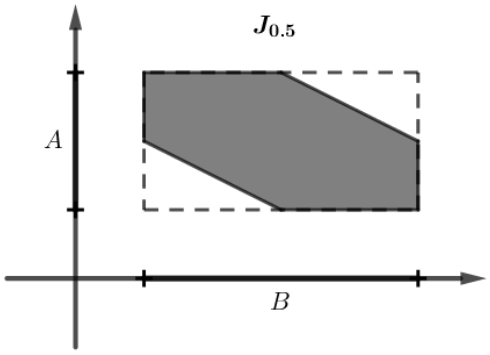


Fig. 2. Graphical representation of the joint possibility distributions $J_{0.5}$ of the intervals A and B.

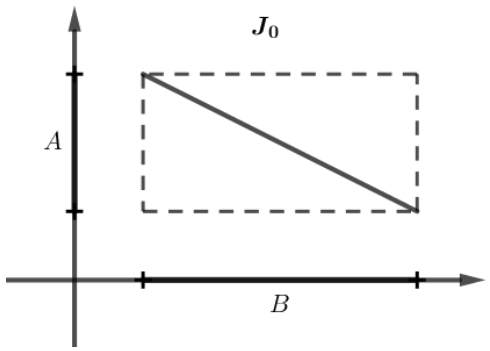


Fig. 3. Graphical representation of the joint possibility distributions J_0 of the intervals A and B.

The gray regions represent the set of all pairs (x_1, x_2), with $x_1 \in A$ and $x_2 \in B$, such that $J_\gamma(x_1, x_2) > 0$. As much the value of γ decreases, the small is the gray region and the higher is the interactivity between A and B.

Note that for $\gamma = 1$ the distribution J_1 is equivalent to the Cartesian product $A \times B$. In this case, from the perspective of fuzzy set theory, the intervals are called non-interactive. For $\gamma = 0$ we obtain the smallest region. In this case the intervals A and B have the highest "level" of interactivity, which is similar to the case of completely correlation [4].

In the next theorem we provide the interactive sum between intervals via joint possibility distribution J_γ, for any value of $\gamma \in [0, 1]$.

Theorem 1. *[13] Let $A = [\underline{a}, \overline{a}]$ and $B = [\underline{b}, \overline{b}]$ be two intervals and $\gamma \in [0, 1]$. Consider $A^t = [\underline{a}^t, \overline{a}^t]$ and $B^t = [\underline{b}^t, \overline{b}^t]$ the translated versions of A and B by their midpoints $a = 0.5(\underline{a}+\overline{a})$ and $b = 0.5(\underline{b}+\overline{b})$, respectively. The interactive sum $C = A +_\gamma B$ for any $\gamma \in [0, 1]$ is given by $C = [\underline{c}, \overline{c}]$, where*

$$\begin{aligned} \underline{c} &= \Big(\underline{a}^t + \overline{b}^t - \gamma(\overline{b}^t - \underline{b}^t)\Big) \wedge \big(\overline{a}^t + \underline{b}^t - \gamma(\overline{a}^t - \underline{a}^t)\big) \wedge \\ &\quad \big(\gamma(\underline{a}^t + \underline{b}^t)\big) + \{a + b\}, \end{aligned}$$

and

$$\begin{aligned} \overline{c} &= \Big(\underline{a}^t + \overline{b}^t + \gamma(\overline{a}^t - \underline{a}^t)\Big) \vee \Big(\overline{a}^t + \underline{b}^t + \gamma(\overline{b}^t - \underline{b}^t)\Big) \vee \\ &\quad \Big(\gamma(\overline{a}^t + \overline{b}^t)\Big) + \{a + b\}. \end{aligned}$$

It is important to observe that in the construction of the interactive sum $+_\gamma$, the intervals are centered on the origin first, then they are computed, and after that the translation is offset by the value $\{a + b\}$. The next example illustrates this calculation.

Example: Let $A = [1, 2]$ and $B = [2, 4]$ be intervals. Translating A and B by their midpoints $a = 1.5$ and $b = 3$, respectively, we obtain $A^t = [-0.5, 0.5]$ and $B^t = [-1, 1]$.

Hence, $C = A +_\gamma B$ is given by

$$\begin{aligned} C &= [-0.5 + 1 - \gamma(1-(-1)) \wedge \\ &\quad 0.5 + (-1) - \gamma(0.5-(-0.5)) \wedge \\ &\quad \gamma(-0.5-1)\ , \\ &\quad -0.5 + 1 + \gamma(0.5-(-0.5)) \vee \\ &\quad 0.5 + (-1) + \gamma(1-(-1)) \vee \\ &\quad \gamma(0.5+1)] + \{1.5+3\} \\ &= [0.5 - 2\gamma \wedge -0.5 - \gamma \wedge -1.5\gamma\ , \\ &\quad 0.5 + \gamma \vee -0.5 + 2\gamma \vee 1.5\gamma] + \{4.5\} \\ &= [-0.5-\gamma, 0.5+\gamma] + \{4.5\} \\ &= [4-\gamma, 5+\gamma]. \end{aligned}$$

Hence,

$$\begin{aligned} A +_0 B &= [4,5] \\ A +_{0.25} B &= [3.75, 5.25] \\ A +_{0.5} B &= [3.5, 5.5] \\ A +_{0.75} B &= [3.25, 5.75] \\ A +_1 B &= [3,6]. \end{aligned}$$

Note that for $\gamma = 1$, we obtain the usual arithmetic sum for intervals. This is a property of the interactive arithmetic associated with J_γ. For more details and properties, the reader can refer to [11].

The interactive difference between intervals is defined by $A -_\gamma B = A +_\gamma (-B) = [\underline{a}, \overline{a}] +_\gamma [-\overline{b}, -\underline{b}]$. This definition gives raise to the interactive derivative for interval-valued functions via J_γ.

Definition 8. *Let $f : (a,b) \to \mathbb{R}$ be an interval-valued function. The function f is J_γ-differentiable at t_0 if the limit*

$$\lim_{h\to 0} \frac{f(t_0+h) -_{\gamma(h)} f(t_0)}{h}$$

exists and it is equal to some element $f'_\gamma(t_0) \in \mathbb{R}$ under the metric d_∞. The number $f'_\gamma(t_0)$ is called J_γ-derivative of f at t_0.

It is important to observe that $\gamma(h)$ depends on h, since for each h we may have a different interactivity between $f(t_0)$ and $f(t_0+h)$. Consequently, $\gamma(h)$ can be view as function in the form of $\gamma : (-h, h) \to [0,1]$.

Considering y such that

$$y = \frac{x(t+h) -_{\gamma(h)} x(t)}{h} \Leftrightarrow yh = x(t+h) -_{\gamma(h)} x(t),$$

for some h, we have that

$$x(t+h) = yh +_{\gamma(h)} x(t).$$

If h is sufficiently small, then y is an approximation for the interactive derivative x'_γ, if it exists. The values that $\gamma(h)$ assume can be modeled by a dataset. For example, Sussner *et al.* [13] considered a chemical degradation problem, where the "level" (that is, the value $\gamma \in [0,1]$) of interactivity has been adjusted by a dataset. Depending on the dataset, different values of γ were obtained.

In the next section we discuss the consistence between the analytical solution via gH-derivative and numerical solution via J_γ-interactive arithmetic. To this end, we consider an extension of Euler's method, that is, the classical arithmetic operations in Euler's method is extended for an arithmetic between interval values.

III. Numerical approximation to the IVP with interval initial value condition

Consider the following IVP

$$\begin{cases} x'_{gH} = f(t,x) \\ x(t_0) = x_0 = [x_0^-, x_0^+] \end{cases}. \tag{6}$$

The analytical solution of (6) can be obtained from a system of classical differential equations [14]. If the gH-derivative is the type I, then we obtain

$$\begin{cases} [x'^-, x'^+] = [f^-, f^+] \\ x(t_0) = [x_0^-, x_0^+] \end{cases} \Rightarrow \begin{cases} x'^- = f^- \\ x'^+ = f^+ \\ x(t_0) = [x_0^-, x_0^+] \end{cases}. \tag{7}$$

If the gH-derivative is the type II, then we obtain

$$\begin{cases} [x'^-, x'^+] = [f^+, f^-] \\ x(t_0) = [x_0^-, x_0^+] \end{cases} \Rightarrow \begin{cases} x'^- = f^+ \\ x'^+ = f^- \\ x(t_0) = [x_0^-, x_0^+] \end{cases}. \tag{8}$$

The purpose here is to provide a numerical approximation for the analytical solution of (6). From Euler's method we have that

$$x_{k+1} = x_k +_{\gamma(h)} hf(t, x_k). \tag{9}$$

Hence, we must prove that x_{k+1} approximates the analytical solution of (6), that is,

$$\lim_{h\to 0} \frac{||x_{k+1} -_{gH} (x_k +_{\gamma(h)} hf(t,x_k))||}{h} = 0$$

Consider a general value of γ and let us compute the right side of Equation (9). We obtain

$$\begin{aligned} [x_{k+1}^-, x_{k+1}^+] &= [x_k^-, x_k^+] +_\gamma [hf_k^-, hf_k^+] \\ &= [(x_k^{-t} + hf_k^{+t} - \gamma h(f_k^{+t} - f_k^{-t})) \wedge \\ &\quad (x_k^{+t} + hf_k^{-t} - \gamma(x_k^{+t} - x_k^{-t})) \wedge \\ &\quad \gamma(x_k^{-t} + hf_k^{-t})\ , \\ &\quad (x_k^{-t} + hf_k^{+t} + \gamma(x_k^{+t} - x_k^{-t})) \vee \\ &\quad (x_k^{+t} + hf_k^{-t} + \gamma h(f_k^{+t} - f_k^{-t})) \vee \\ &\quad \gamma(x_k^{+t} + hf_k^{+t})] + \{x+f\}, \end{aligned} \tag{10}$$

where $x = 0.5(x_k^+ + x_k^-)$ and $f = 0.5h(f_k^+ + f_k^-)$.

The task to determinate the left and right sides of the interval (10), depends on the values that f_k^- and f_k^+ assume. In particular for $\gamma = 1$, we obtain:

$$\begin{aligned} [x_{k+1}^-, x_{k+1}^+] &\overset{\text{for } \gamma=1}{=} [x_k^{-t} + hf_k^{-t}, x_k^{+t} + hf_k^{+t}] + \{x+f\} \\ &= [x_k^- + hf_k^-, x_k^+ + hf_k^+] \end{aligned}$$

We observe that numerical solution given by $\gamma = 1$ is the approximation for the analytical solution via gH-derivative of type I.

For $\gamma = 0$, we obtain:

$$\begin{aligned}[x_{k+1}^-, x_{k+1}^+] &\overset{\text{for } \gamma=0}{=} [(x_k^{-t} + hf_k^{+t}) \wedge (x_k^{+t} + hf_k^{-t}) \wedge 0,\\ &\quad (x_k^{-t} + hf_k^{+t}) \vee (x_k^{+t} + hf_k^{-t}) \vee 0]\\ &+ \{x + f\}\\ &= [(x_k^- + hf_k^+) \wedge (x_k^+ + hf_k^-) \wedge x + f,\\ &\quad (x_k^- + hf_k^+) \vee (x_k^+ + hf_k^-) \vee x + f]\end{aligned}$$

Since x and f are the midpoints of the intervals $[x_k^-, x_k^+]$ and $[hf_k^+, hf_k^-]$, respectively, it follows that

$$\begin{aligned}[x_{k+1}^-, x_{k+1}^+] &\overset{\text{for } \gamma=0}{=} [(x_k^- + hf_k^+) \wedge (x_k^+ + hf_k^-),\\ &\quad (x_k^- + hf_k^+) \vee (x_k^+ + hf_k^-)].\end{aligned}$$

Hence, the numerical solution given by $\gamma = 0$ is the approximation for the analytical solution via gH-derivative of type II. These calculations give raise to the following theorem.

Theorem 2. *Let x be a solution of* (6) *in the sense of gH-differentiability. Thus:*

- *Euler's method, with the J_1-interactive arithmetic, approximates the analytical solution of gH-derivative of type I;*
- *Euler's method, with the J_0-interactive arithmetic, approximates the analytical solution of gH-derivative of type II.*

In the next section we illustrate the result of Theorem 2 in the decay Malthusian model under the gH-differentiability.

IV. The Fuzzy Malthusian Model

Consider the decay Malthusian model, with interval initial condition given by

$$\begin{cases} x'_{gH} = -\lambda x, \quad \lambda > 0 \\ x(0) = [x_0^-, x_0^+] \end{cases} \tag{11}$$

For this interval initial value problem we consider two cases:

1) Hukuhara derivative ((I)-gH-derivative)
2) Generalized Hukuhara derivative ((II)-gH-derivative)

A. (I)-gH-derivative

Considering the (I)-gH-derivative in the Malthusian model, we obtain the following associated system of classical differential equations.

$$\begin{cases} x'^-(t) = -\lambda x^+(t), \quad \lambda > 0 \\ x'^+(t) = -\lambda x^-(t) \\ x^-(0) = x_0^- \\ x^+(0) = x_0^+ \end{cases}. \tag{12}$$

From Euler's method with interactive arithmetic via J_γ for intervals, given by (10), we have

$$\begin{aligned}[x_{k+1}^-, x_{k+1}^+] &= [(x_k^{-t} - h\lambda x_k^{-t} - \gamma h(-\lambda x_k^{-t} + \lambda x_k^{+t})) \wedge\\ &\quad (x_k^{+t} - h\lambda x_k^{+t} - \gamma(-x_k^{-t} + x_k^{+t})) \wedge\\ &\quad \gamma(x_k^{-t} - h\lambda x_k^{+t}),\\ &\quad (x_k^{-t} - h\lambda x_k^{-t} + \gamma(x_k^{+t} - x_k^{-t})) \vee\\ &\quad (x_k^{+t} - h\lambda x_k^{+t} + \gamma h(-\lambda x_k^{-t} + \lambda x_k^{+t})) \vee\\ &\quad \gamma((x_k^{+t} - h\lambda x_k^{-t}))]\\ &+ \left\{\frac{(1-\lambda h)(x_k^- + x_k^+)}{2}\right\}\end{aligned}$$

Note that $(x_k^{-t} - h\lambda x_k^{-t} - \gamma h(-\lambda x_k^{-t} + \lambda x_k^{+t})) \leq (x_k^{+t} - h\lambda x_k^{+t} - \gamma(-x_k^{-t} + x_k^{+t}))$ if, and only if $x_k^{-t}((1-\gamma)(1-h\lambda)) \leq x_k^{+t}((1-\gamma)(1-h\lambda))$. Since we are considering small values of $h > 0$ and $\lambda > 0$, this inequality holds true.

Also, $(x_k^{-t} - h\lambda x_k^{-t} - \gamma h(-\lambda x_k^{-t} + \lambda x_k^{+t})) \leq \gamma(x_k^{-t} - h\lambda x_k^{+t})$ if, and only if, $x_k^{-t}((1-\gamma)(1-h\lambda)) \leq 0$. Since $x_k^{-t} \leq 0$ and $(1-\gamma)(1-h\lambda) \geq 0$, for small values of $h > 0$ and $\lambda > 0$, this inequality holds true.

Consequently, we obtain

$$x_{k+1}^{-t} = (x_k^{-t} - h\lambda x_k^{-t} - \gamma h(-\lambda x_k^{-t} + \lambda x_k^{+t})).$$

Similarly, $(x_k^{-t} - h\lambda x_k^{-t} + \gamma(x_k^{+t} - x_k^{-t})) \leq (x_k^{+t} - h\lambda x_k^{+t} + \gamma h(-\lambda x_k^{-t} + \lambda x_k^{+t}))$ if, and only if, $x_k^{-t}((1-\gamma)(1-h\lambda)) \leq x_k^{+t}((1-\gamma)(1-h\lambda))$, which occurs. On the other hand, $(x_k^{+t} - h\lambda x_k^{+t} + \gamma h(-\lambda x_k^{-t} + \lambda x_k^{+t})) \geq \gamma((x_k^{+t} - h\lambda x_k^{-t}))$ if, and only if, $x_k^{+t}((1-\gamma)(1-h\lambda)) \geq 0$. Since $x_k^{+t} \geq 0$ and $(1-\gamma)(1-h\lambda) \geq 0$, for small values of $h > 0$ and $\lambda > 0$, this inequality holds true.

Consequently, we obtain

$$x_{k+1}^{+t} = (x_k^{+t} - h\lambda x_k^{+t} + \gamma h(-\lambda x_k^{-t} + \lambda x_k^{+t})).$$

Therefore,

$$\begin{aligned}[x_{k+1}^-, x_{k+1}^+] &= [(x_k^{-t} - h\lambda x_k^{-t} - \gamma h(-\lambda x_k^{-t} + \lambda x_k^{+t})),\\ &\quad (x_k^{+t} - h\lambda x_k^{+t} + \gamma h(-\lambda x_k^{-t} + \lambda x_k^{+t}))]\\ &+ \left\{\frac{(1-\lambda h)(x_k^- + x_k^+)}{2}\right\}.\end{aligned}$$

In order to ensure that

$$\lim_{h\to 0} \frac{||x_{k+1} -_{(I)-gH} (x_k +_\gamma hf(t, x_k))||}{h} = 0,$$

we must have

$$(x_k^{-t} - h\lambda x_k^{-t} - \gamma h(-\lambda x_k^{-t} + \lambda x_k^{+t})) \to x_k^{-t} - h\lambda x_k^{+t}$$

and

$$(x_k^{+t} - h\lambda x_k^{+t} + \gamma h(-\lambda x_k^{-t} + \lambda x_k^{+t})) \to x_k^{+t} - h\lambda x_k^{-t},$$

which only occurs when $\gamma \to 1$. Consequently, for $\gamma = 1$ we obtain

$$\begin{aligned}[x_{k+1}^-, x_{k+1}^+] &= [x_k^{-t} - h\lambda x_k^{+t}, x_k^{+t} - h\lambda x_k^{-t}]\\ &+ \left\{\frac{(1-\lambda h)(x_k^- + x_k^+)}{2}\right\}\\ &= [x_k^- - h\lambda x_k^+, x_k^+ - h\lambda x_k^-].\end{aligned}$$

The above expression is the Euler's method for (12). This means that the numerical solution via interactive arithmetic, for $\gamma = 1$, is compatible with the analytical solution of the first form of gH-derivative, corroborating the result stated in Theorem 2.

B. (II)-gH-derivative

For gH-derivative of type II, we have the following system

$$\begin{cases} x'^{-}(t) = -\lambda x^{-}(t), \quad \lambda > 0 \\ x'^{+}(t) = -\lambda x^{+}(t), \\ x^{-}(0) = x_0^{-} \\ x^{+}(0) = x_0^{+} \end{cases} \tag{13}$$

In order to ensure that

$$\lim_{h \to 0} \frac{||x_{k+1} -_{(II)-gH} (x_k +_\gamma hf(t, x_k))||}{h} = 0,$$

we must have

$$(x_k^{-t} - h\lambda x_k^{-t} - \gamma h(-\lambda x_k^{-t} + \lambda x_k^{+t})) \to x_k^{-t} - h\lambda x_k^{-t}$$

and

$$(x_k^{+t} - h\lambda x_k^{+t} + \gamma h(-\lambda x_k^{-t} + \lambda x_k^{+t})) \to x_k^{+t} - h\lambda x_k^{+t},$$

which only occurs when $\gamma \to 0$. Consequently, for $\gamma = 0$ we obtain

$$\begin{aligned} [x_{k+1}^{-}, x_{k+1}^{+}] &= [x_k^{-t} - h\lambda x_k^{-t}, x_k^{+t} - h\lambda x_k^{+t}] \\ &+ \left\{ \frac{(1 - \lambda h)(x_k^{-} + x_k^{+})}{2} \right\} \\ &= [x_k^{-} - h\lambda x_k^{-}, x_k^{+} - h\lambda x_k^{+}]. \end{aligned}$$

The above expression is the Euler's method for (13). This means that the numerical solution via interactive arithmetic, for $\gamma = 0$, is compatible with the analytical solution of the second form of gH-derivative, corroborating the result stated in Theorem 2.

Next we present a brief summary of the numerical approximations provided in this article.

TABLE I
SUMMARY OF NUMERICAL APPROXIMATIONS BETWEEN EULER'S METHOD WITH INTERACTIVE ARITHMETIC AND ANALYTICAL SOLUTIONS VIA gH-DERIVATIVE

Euler's method	Value of γ	Approximated solution
$x_k +_{\gamma(h)} hf(t, x_k)$	$\gamma(h) \overset{h \to 0}{\to} 1$	(I)-gH-derivative
$x_k +_{\gamma(h)} hf(t, x_k)$	$\gamma(h) \overset{h \to 0}{\to} 0$	(II)-gH-derivative

V. FINAL REMARKS

In this article we study numerical methods to approximate analytical solutions for Initial Value Problems, with interval initial conditions. Here we focus on Euler's method, whose classical arithmetic operations are adapted for interval values. To this end, we consider an arithmetic employed by Esmi *et al.* [11], which is obtained from the sup-J extension principle, where J, for this case, refers to a parametrized family of joint possibility distributions J_γ between intervals.

We showed that once considered the interactive arithmetic via J_γ, where $\gamma \in [0, 1]$, the obtained numerical solution approximates the analytical solution referring to the approach via gH-derivative. That is, if γ goes to 1, then the numerical solution via Euler's method approximates the analytical solution of gH-derivative of type I (*i.e*, Hukuhara derivative). On the other hand, if γ goes to 0, then the numerical solution approximates the analytical solution of gH-derivative of type II (see TABLE I). This result is presented in Theorem 2.

For a better illustration of the results, we provide more detailed calculations for a decay Malthusian model, which analytical solution is given by an exponential function.

It is important to observe that we prove this connection for an interval initial value problem. As a future work, this result will be extended for a fuzzy initial value problem. Also, we intend to study an error analysis for such approximations.

REFERENCES

[1] L. Edelstein-Keshet, Mathematical Models in Biology, 1st ed., vol. 1. Society for Industrial and Applied Mathematic, 2006.

[2] O. Kaleva, "Fuzzy differential equations," Fuzzy Sets and Systems, vol. 24, Issue 3, pp. 301-317, December 1987.

[3] B. Bede, Mathematics of Fuzzy Sets and Fuzzy Logic, Springer, 2012.

[4] L. C. de Barros, and F. S. Pedro, "Fuzzy differential equations with interactive derivative," Fuzzy Sets and Systems, vol. 309, pp. 64-80, February 2017.

[5] E. Esmi, D. E. Sánchez, V. F. Wasques and L. C. de Barros, "Solutions of higher order linear fuzzy differential equations with interactive fuzzy values," Fuzzy Sets and Systems, vol. 419, pp. 122-140, August 2021.

[6] M. T. Mizukoshi, L. C. Barros, Y. Chalco-Canoc, H. Román-Flores, and R.C.Bassanezi, "Fuzzy differential equations and the extension principle," Information Sciences, vol. 177, Issue 17, pp. 3627-3635, September 2007.

[7] V. F. Wasques, E. Esmi, L. C. Barros, and P. Sussner, "The generalized fuzzy derivative is interactive," Information Sciences, vol. 519, pp. 93-109, May 2020.

[8] D. Dubois, and H. Prade, Fuzzy Sets and Systems: Theory and Applications, 1st ed., Mathematics in Science and Engineering, Academic Press, 1980.

[9] T. Allahviranloo and S. Salahshour, Advances in Fuzzy Integral and Differential Equations, Springer, Cham, 1st ed, 2022.

[10] V. F. Wasques, E. Esmi, L. C. Barros, and P. Sussner, "Numerical Solution for Fuzzy Initial Value Problems via Interactive Arithmetic: Application to Chemical Reactions," International Journal of Computational Intelligence Systems, vol. 13, Issue 1, pp. 1517 - 1529, September 2020.

[11] E. Esmi , V. F. Wasques, and L. C. de Barros, "Addition and subtraction of interactive fuzzy numbers via family of joint possibility distributions," Fuzzy Sets and Systems, vol. 424, pp. 105-131, November 2021.

[12] L. A. Zadeh, "Fuzzy sets," Information and Control, vol. 8, Issue 3, pp. 338-353, June 1965.

[13] P. Sussner, V. Wasques, J. C. Fernandes, and E. Esmi, A Numerical Approach For Determining Interactivity: A Case Study in Chemical Degradation, Joint Proc. of the 12th Conf. of the European Society for Fuzzy Logic and Technology (EUSFLAT), 2021.

[14] T. Allahviranloo, and W. Pedrycz, Soft numerical computing in uncertain dynamic systems, Elsevier, Academic Press, 2020.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

On P-torsion EQ-modules and P-cyclic EQ-modules

1st Batoul Ganji Saffar

1st Department of Mathematics, Faculty of Mathematical Sciences, Alzahra University, Tehran, Iran, bganji@alzahra.ac.ir

***Abstract*— We applied the module theory the same as ring theory on EQ-algebras. With the concept of the maximal ideal on EQ-algebras, we introduce notions of P-torsion EQ-modules and P-cyclic EQ-modules, for each maximal ideal of E. Then, we gain relations between these notions and multiplicative EQ modules.**

***Index Terms*— EQ-algebras, EQ-modules, maximal ideal, $\mathfrak{P}$-torsion EQ-modules, $\mathfrak{P}$-cyclic EQ-modules, multiplicative EQ-modules.**

I. Preliminaries

A type of logical algebras named EQ-algebra was proposed by Vilém Novák in [6]. This algebra generalizes residuated lattice and said as an algebraic structure of truth value for fuzzy type theory. The operations of EQ-algebra are multiplication, meet and fuzzy equality $(\otimes_{\Im}, \wedge_{\Im}, \sim_{\Im})$. Also implication $(\multimap)$ is defined by $\zeta \multimap \varpi = (\zeta \wedge_{\Im} \varpi) \sim_{\Im} \zeta$. In fact, unlike Hájek [4], we can gain non-commutatively without to defined two types of implications.
We know each module is constructed by an action of a ring on an abelian group. In some logical algebraic structures analog MV-algebras and BCK-algebras [1], they used the notions of ring theory, verified the action of these algebras on groups. They investigated " modules" and their properties on these algebras. We know that, these algebras are special cases of EQ-algebras, thus we trying to construct a same theory on EQ-algebras. In this paper, by using of the concepts of the maximal ideal on EQ-algebras, we define concepts of P-torsion EQ-modules and P-cyclic EQ-modules, for each maximal ideal of $\Im_E$ and gain interesting relation between these concepts and multiplicative EQ-modules.

II. Introduction

An algebraic structure $\Im_E = (E_{\Im}, \wedge_{\Im}, \otimes_{\Im}, \sim_{\Im}, 1)$ of type (2, 2, 2, 0), is called an *EQ-algebra*, if for each elements $\zeta, \varpi, \varrho, \varsigma \in E_{\Im}$, we have:
$(E_{\Im}1)$ $\langle E_{\Im}, \wedge_{\Im}, 1\rangle$ is a commutative idempotent monoid (i.e. $\wedge_{\Im}$-semilattice with top element 1);
$(E_{\Im}2)$ $\langle E_{\Im}, \otimes_{\Im}, 1\rangle$ is a monoid and $\otimes_{\Im}$ is isotone w.r.t. " $\leq_{\Im}$ "
(where $\zeta \leq_{\Im} \varpi$ i.e. $\zeta \wedge_{\Im} \varpi = \zeta$);
$(E_{\Im}3)$ $\zeta \sim_{\Im} \zeta = 1$;
$(E_{\Im}4)$ $((\zeta \wedge_{\Im} \varpi) \sim_{\Im} \varrho) \otimes_{\Im} (\varsigma \sim_{\Im} \zeta) \leq_{\Im} \varrho \sim_{\Im} (\varsigma \wedge_{\Im} \varpi)$;
$(E_{\Im}5)$ $(\zeta \sim_{\Im} \varpi) \otimes_{\Im} (\varrho \sim_{\Im} \varsigma) \leq_{\Im} (\zeta \sim_{\Im} \varrho) \sim_{\Im} (\varpi \sim_{\Im} \varsigma)$;
$(E_{\Im}6)$ $(\zeta \wedge_{\Im} \varpi \wedge_{\Im} \varrho) \sim_{\Im} \zeta \leq_{\Im} (\zeta \wedge_{\Im} \varpi) \sim_{\Im} \zeta$;
$(E_{\Im}7)$ $\zeta \otimes_{\Im} \varpi \leq_{\Im} \zeta \sim_{\Im} \varpi$.
Let $\Im_E = (E_{\Im}, \wedge_{\Im}, \otimes_{\Im}, \sim_{\Im}, 1)$ be an EQ-algebra. Then $\Im_E$ is called:
(ι) *good* if $\zeta \sim_{\Im} 1 = \zeta = 1 \sim_{\Im} \zeta$, for all $\zeta \in E_{\Im}$;
$(\iota\iota)$ *involutive(IEQ-algebra)* if $E_{\Im}$ contains bottom element "0" and $\neg\neg\zeta = \zeta$, for all $\zeta \in E_{\Im}$ ($\neg\zeta = \zeta \sim_{\Im} 0$).

Proposition 1. *[7] Each IEQ-algebra is a good EQ-algebra.*

Suppose that $\Im_E$ is an EQ-algebra with bottom element "0". Define binary operation $\boxplus_{\Im}$ on $\Im_E$ as follows:

$$\zeta \boxplus_{\Im} \varpi = \neg\zeta \multimap_{\Im} \varpi$$

A nonempty subset $\mathfrak{I}$ of an EQ-algebra $\Im_E$ is called *ideal*, if (ι) $\zeta \boxplus_{\Im} \varpi \in \mathfrak{I}$, for all $\zeta, \varpi \in \mathfrak{I}$ and (ii) $\zeta \leq_{\Im} \varpi$ and $\varpi \in \mathfrak{I}$, then $\zeta \in \mathfrak{I}$.

Let $\mathfrak{A}$ be a nonempty set of EQ-algebra $\Im_E$. Then smallest ideal of $\Im_E$ including $\mathfrak{A}$ is called to be the ideal generated by $\mathfrak{A}$ and is denoted by $< \mathfrak{A} >$.

Theorem 1. *[5] For each nonempty set $\mathfrak{T}$ of an EQ-algebra $\Im_E$, we have* $< \mathfrak{T} >= \{\varsigma \in E_{\Im} \mid \varsigma \leq (...((\varrho_1 \boxplus_{\Im} \varrho_2) \boxplus_{\Im} \varrho_3)... \boxplus_{\Im} \varrho_\nu), \varrho_i \in \mathfrak{T}, i = 1, 2, ..., \nu\}$.

Proposition 2. *[5] Let $\mathfrak{I}$ be an ideal of $\Im_E$ and $\iota \in E_{\Im}$. Then* $< \mathfrak{I} \cup_{\Im} \{\iota\} >= \{\varrho \in E_{\Im} \mid \exists \nu \in \mathbb{N}, \mathfrak{q} \in \mathfrak{I}, s.t.\ \varrho \leq_{\Im} (\nu\iota) \boxplus_{\Im} \mathfrak{q}\}$.

Definition 1. *[3] Let $\Im_E = (E_{\Im}, \wedge_{\Im}, \otimes_{\Im}, \sim_{\Im}, 1)$ be an EQ-algebra, $\mathfrak{M}_{\Im}$ be an abelian group under " $+$ " and $(\zeta, \mu) \multimap_{\Im} \zeta.\mu$ be a mapping of $E_{\Im} \times \mathfrak{M}_{\Im} \multimap_{\Im} \mathfrak{M}_{\Im}$ such that, for all $\mu \in \mathfrak{M}_{\Im}$ and $\zeta, \varpi \in E_{\Im}$:*
$(E_{\Im}\mathfrak{M}1)$ $1.\mu = \mu$,
$(E_{\Im}\mathfrak{M}2)$ $(\zeta \wedge_{\Im} \varpi).\mu = \zeta.(\varpi.\mu)$,
$(E_{\Im}\mathfrak{M}3)$ $\zeta.(\mu_1 + \mu_2) = \zeta.\mu_1 + \zeta.\mu_2$,
$(E_{\Im}\mathfrak{M}4)$ $(\zeta \sim_{\Im} \varpi).\mu = \mu - (\zeta.\mu + \varpi.\mu)$ *if $\zeta \neq \varpi$. where the subtraction "-"is in fact the group subtraction in $\mathfrak{M}_{\Im}$. Then $\mathfrak{M}_{\Im}$ is called an EQ-module.*

We write $\zeta\mu$ instead of $\zeta.\mu$, for short.

Proposition 3. *[3] Let $\Im_E = (E_\Im, \wedge_\Im, \otimes_\Im, \sim_\Im, 0, 1)$ be an IEQ-algebra such that $\zeta \multimap_\Im \varpi = \neg\zeta \vee_\Im \varpi$ and $\zeta \vee_\Im \neg\zeta = 1$, for all $\zeta, \varpi \in E_\Im$. Then by defining a binary operation $\zeta + \varpi = (\zeta \multimap_\Im \varpi) \wedge_\Im (\varpi \multimap_\Im \zeta)$ on $\mathfrak{M}_\Im = E_\Im$, $(\mathfrak{M}_\Im, +)$ is an abelian group. Also, if $+ := \sim_\Im$ and $. : E_\Im \times \mathfrak{M}_\Im \multimap_\Im \mathfrak{M}_\Im$ is defined by $.(\zeta, \mu) = \zeta.\mu = \zeta \multimap_\Im \mu$, then $\mathfrak{M}_\Im$ is an EQ-module.*

Proposition 4. *[3] Let $\Im_E$ be an EQ-algebra with bottom element "0" and $\mathfrak{M}_\Im$ be an EQ-module. Then, for all $\zeta, \varpi \in E_\Im$ and $\mu, \nu \in \mathfrak{M}_\Im$ the following conditions hold:*
(ι) $\zeta 0 = 0$;
$(\iota\iota)$ $\zeta(-\mu) = -(\zeta\mu)$;
$(\iota\iota\iota)$ *If $0\mu = 0$, then $(\neg\zeta)\mu = \mu - \zeta\mu$;*
$(\iota\nu)$ $\zeta - \varpi\mu) = -(\zeta(\varpi\mu))$.

Definition 2. *[3] Let $\mathfrak{M}_\Im$ be an EQ-module and $\emptyset \neq \mathfrak{N}_\Im \subseteq \mathfrak{M}_\Im$. Then $\mathfrak{N}_\Im$ is called a* submodule *of $\mathfrak{M}_\Im$ if for all $\nu, \nu' \in \mathfrak{N}_\Im$ and $\zeta \in E_\Im$: (ι) $\nu - \nu' \in \mathfrak{N}_\Im$, $(\iota\iota)$ $\zeta\nu \in \mathfrak{N}_\Im$.*

Example 1. *[3] Let $\mathfrak{M}_\Im$ be an EQ-module. Then $E_\Im\mu = \{\zeta\mu \mid \zeta \in E_\Im\}$ is a submodule of $\mathfrak{M}_\Im$, for all $\mu \in \mathfrak{M}_\Im$.*

Definition 3. *[3] Let $\mathfrak{M}_\Im$ be an EQ-module and $\mathfrak{N}_\Im$ be a submodule of $\mathfrak{M}_\Im$. We define $(\mathfrak{N}_\Im :_{E_\Im} \mathfrak{M}_\Im)$ as follows:*

$$(\mathfrak{N}_\Im :_{E_\Im} \mathfrak{M}_\Im) = \{\zeta \in E_\Im \mid \zeta\mathfrak{M}_\Im \subseteq \mathfrak{N}_\Im\}.$$

We denote annihilator of N by $Ann_E(N)$, which is defined by

$$Ann_{E_\Im}(\mathfrak{N}_\Im) = (0 :_{E_\Im} \mathfrak{N}_\Im) = \{\zeta \in E_\Im \mid \zeta\mathfrak{N}_\Im = 0\}.$$

Proposition 5. *[3] Let $\Im_E$ be an IEQ-algebra, $\mathfrak{M}_\Im$ be an EQ-module such that $0\mu = 0$, for all $\mu \in \mathfrak{M}_\Im$, and $\mathfrak{N}_\Im$ be a submodule of $\mathfrak{M}_\Im$. Then $(\mathfrak{N}_\Im :_{E_\Im} \mathfrak{M}_\Im)$ is an ideal of $\Im_E$.*

Definition 4. *[3] Let $\Im_E$ be an EQ-algebra and $\mathfrak{M}_\Im$ be an EQ-module. Then $\mathfrak{M}_\Im$ is called a* multiplication EQ-module*, if for every submodule $\mathfrak{N}_\Im$, there exists an ideal $\mathfrak{I}$ of $\Im_E$ such that $\mathfrak{N}_\Im = \mathfrak{I}\mathfrak{M}_\Im$.*

Proposition 6. *[3] Let $\mathfrak{M}_\Im$ be an EQ-module. Then $\mathfrak{M}_\Im$ is a multiplication EQ-module if and only if $\mathfrak{N}_\Im = (\mathfrak{N}_\Im :_{E_\Im} \mathfrak{M}_\Im)\mathfrak{M}_\Im$, for every submodule $\mathfrak{N}_\Im$ of $\mathfrak{M}_\Im$.*

Proposition 7. *[3] Every cyclic EQ-module is a multiplication EQ-module.*

Proposition 8. *[3] Homomorphic image of every multiplication module is a multiplication module.*

III. $\mathfrak{P}$-TORTION EQ-MODULES AND $\mathfrak{P}$-CYCLIC EQ-MODULES

In this section, we define the concepts of the maximal ideal on EQ-algebras, then we introduce concepts of $\mathfrak{P}$-torsion EQ-modules and $\mathfrak{P}$-cyclic EQ-modules, for each maximal ideal of $\Im_E$ and gain interesting relation between these concepts and multiplicative EQ-modules.

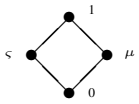
Definition 5. *Let $\Im_E$ be a lattice ordered EQ-algebra and $\mathfrak{I}$ be an proper ideal of $\Im_E$. We called $\mathfrak{I}$ a* maximal *ideal of $\Im_E$, if it is not properly contained in any other ideal of $\Im_E$.*

Example 2. *Let $E_\Im = \{0, \varsigma, \mu, 1\}$. Define $\wedge_\Im, \otimes_\Im, \sim_\Im$ as follows:*

$\otimes_\Im$	0	ς	μ	1
0	0	0	0	0
ς	0	ς	0	ς
μ	0	0	μ	μ
1	0	ς	μ	1

$\sim_\Im$	0	ς	μ	1
0	1	μ	ς	0
ς	μ	1	0	ς
μ	ς	0	1	μ
1	0	ς	μ	1

$\multimap_\Im$	0	ς	μ	1
0	1	1	1	1
ς	μ	1	μ	1
μ	ς	ς	1	1
1	0	ς	μ	1



Then $\Im_E = (E_\Im, \wedge_\Im, \otimes_\Im, \sim_\Im, 0, 1)$ is an IEQ-algebra such that $\neg\zeta \vee_\Im \zeta = 1$ and $\zeta \multimap_\Im \varpi = \neg\zeta \vee_\Im \varpi$, for all $\zeta, \varpi \in E_\Im$ and by definition binary operation $\zeta + \varpi = (\zeta \multimap_\Im \varpi) \wedge_\Im (\varpi \multimap_\Im \zeta)$ on $\mathfrak{M}_\Im = E_\Im$, $(\mathfrak{M}_\Im, +)$ is an abelian group. Also, if $+ = \sim_\Im$ and define $. : E_\Im \times \mathfrak{M}_\Im \multimap_\Im \mathfrak{M}_\Im$ by $.(\zeta, \mu) = \zeta\mu = \mu$, if $\zeta = 1$ and $.(\zeta, \mu) = \zeta\mu = 0$, if $\zeta \neq 1$. Then $\mathfrak{M}_\Im$ is an EQ-module.

Example 3. *In Example 2, let $\mathfrak{P} = \{0, \varsigma\}$. Since $\mu \boxplus_\Im \varsigma = \neg\mu \multimap_\Im a = 1$ and $1 \boxplus_\Im 0 = \neg 1 \multimap_\Im 0 = 1$, then $\mathfrak{P}$ is a maximal ideals of $\Im_E$.*

Proposition 9. *Let $\Im_E$ be an EQ-algebra with bottom element "0" and $\mathfrak{I}$ be an proper ideal of $\Im_E$. Then there exists maximal ideal $\mathfrak{L}$ of $\Im_E$ such that $\mathfrak{I} \subseteq \mathfrak{L}$.*

Proof: By Zorn's Lemma, the proof is clear. ■

Corollary 1. *Each EQ-algebra with bottom element "0" has a maximal ideal.*

Proposition 10. *Let $\Im_E$ be an EQ-algebra with bottom element "0" and $\mathfrak{L}$ be a maximal ideal of $\Im_E$. Then $\zeta \in \mathfrak{L}$ if and only if $\neg(\nu\zeta) \in E_\Im - \mathfrak{L}$, for all $\nu \in \mathbb{N}$.*

Proof: Let $\zeta \in \mathfrak{L}$ and there exists $\nu \in \mathbb{N}$ such that $\neg(\nu\zeta) \notin E_\Im - \mathfrak{L}$. Then $\neg(\nu\zeta) \in \mathfrak{L}$. Since $\nu\zeta = \zeta \otimes_\Im \zeta \otimes_\Im ... \otimes_\Im \zeta \leq_\Im \zeta$, we have $\neg\zeta \leq_\Im \neg(\nu\zeta)$ and so $\neg\zeta \in \mathfrak{L}$. Hence $0 = \zeta \wedge_\Im \neg\zeta \in \mathfrak{L}$, which is a contradiction.
Conversely, let $\neg(\nu\zeta) \notin \mathfrak{L}$, for all $\nu \in \mathbb{N}$ and $\zeta \notin \mathfrak{L}$. Then $\mathfrak{L} \subsetneq < \mathfrak{L} \cup_\Im \{\zeta\} >$ and so $< \mathfrak{L} \cup_\Im \{\zeta\} >= E_\Im$. Hence $1 \in < \mathfrak{L} \cup_\Im \{\zeta\} >$. Thus there exist $\varsigma \in \mathfrak{L}$ and $\nu \in \mathbb{N}$ such that $1 \leq_\Im \nu\zeta \boxplus_\Im \varsigma$ or $1 = \nu\zeta \boxplus_\Im \varsigma$, i.e. $\neg(\nu\zeta) \multimap_\Im \varsigma = 1$ or $\neg(\nu\zeta) \leq_\Im \varsigma$. Since $\varsigma \in \mathfrak{L}$, we get $\neg(\nu\zeta) \in \mathfrak{L}$, which is a contradiction. ■

Definition 6. *Let $\Im_E$ be an EQ-algebra with bottom element "0", $\mathfrak{M}_\Im$ be an EQ-module such that $0\mu = 0$, for all $\mu \in$*

$\mathfrak{M}_\Im$ *and* $\mathfrak{P} \in Max(\mathcal{E}_\Im)$. *Then* $\mathfrak{M}_\Im$ *is called* $\mathfrak{P}$-torsion EQ-module, *if for each* $\mu \in \mathfrak{M}_\Im$, *there exists* $\varrho \in \mathfrak{P}$ *such that* $\neg\varrho\mu = 1$.

Example 4. *Let* $E_\Im = \{0, \varsigma, \mu, \iota, \chi, 1\}$ *be a chain such that* $0 < \varsigma < \mu < \iota < \chi < 1$. *Define the operations* $\wedge_\Im, \otimes_\Im$ *and* $\sim_\Im$ *on* $E_\Im$ *as follows:*

$\otimes_\Im$	0	ς	μ	ι	χ	1
0	0	0	0	0	0	0
ς	0	0	0	0	0	ς
μ	0	0	0	0	ς	μ
ι	0	0	0	ς	ς	ι
χ	0	0	ς	ς	ς	χ
1	0	ς	μ	ι	χ	1

$\sim_\Im$	0	ς	μ	ι	χ	1
0	1	ι	μ	ς	0	0
ς	ι	1	μ	ς	ς	ς
μ	μ	μ	1	μ	μ	μ
ι	ς	ς	μ	1	ι	ι
χ	0	ς	μ	ι	1	χ
1	0	ς	μ	ι	χ	1

$\multimap_\Im$	0	ς	μ	ι	χ	1
0	1	1	1	1	1	1
ς	ι	1	1	1	1	1
μ	μ	μ	1	1	1	1
ι	ς	ς	μ	1	1	1
χ	0	ς	μ	ι	1	1
1	0	ς	μ	ι	χ	1

$$\zeta \wedge_\Im \varpi = \min\{\zeta, \varpi\}.$$

Then $\Im_E = (E_\Im, \wedge_\Im, \otimes_\Im, \sim_\Im, 1)$ *is an EQ-algebra. Define* $. : E_\Im \times \mathbb{Z} \multimap_\Im \mathbb{Z}$ *by:*

$$.(\zeta, \nu) = \zeta\nu = \begin{cases} \nu, & \zeta \neq 0 \\ 0, & \zeta = 0 \end{cases}$$

We can easily show that $\mathbb{Z}$ *is an EQ-module. Let* $\mathfrak{P} = \{0, \varsigma\}$. *Since* $\mu \boxplus_\Im \mu = \neg\mu \multimap_\Im \mu = 11$, $\iota \boxplus_\Im \iota = \neg\iota \multimap_\Im \iota = 1$ *and* $1 \boxplus_\Im 0 = \neg 1 \multimap_\Im 0 = 1$, *then* $\mathfrak{P}$ *is a maximal ideal of* $\Im_E$. *For all* $\nu \in \mathbb{Z}$, *we have* $0 \in \mathfrak{P}$ *and* $\neg\varrho\nu = 1\nu = \nu \neq 0$. *Also, for* $0 \neq \nu \in \mathbb{Z}$ *we have* $\neg\varsigma\nu = \iota\nu = \nu \neq 0$. *Therefore,* $\mathbb{Z}$ *is not* $\mathfrak{P}$*-torsion EQ-module.*

Example 5. *In Example 2,* $\mathfrak{I} = \{0, \varsigma\}$ *is a maximal ideal of* $\Im_E$. *We can see that* $\mathfrak{M}_\Im$ *is a* $\mathfrak{P}$*-torsion EQ-module, because for all* $\mu \in \mathfrak{M}_\Im$ *there is a* $\varrho \in \mathfrak{P}$ *such that* $\neg\varrho\mu = \neg\varrho \multimap_\Im \mu = 0$.

Definition 7. *Let* $\Im_E$ *be an EQ-algebra with bottom element "20",* $\mathfrak{M}_\Im$ *be an EQ-module such that* $0\mu = 0$, *for all* $\mu \in \mathfrak{M}_\Im$ *and* $\mathfrak{P} \in Max(\mathcal{E}_\Im)$. *Then* $\mathfrak{M}_\Im$ *is called* $\mathfrak{P}$-cyclic EQ-module, *if there exist* $\mu \in \mathfrak{M}_\Im$ *and* $\varrho \in \mathfrak{P}$ *such that* $\neg\varrho\mathfrak{M}_\Im \subseteq E_\Im\mu$.

Example 6. *In Example 4, let* $\mathfrak{P} = \{0, \varsigma\}$. *Then* $\mathfrak{P}$ *is a maximal ideal of* $\Im_E$ *and* $\mathbb{Z}$ *is a* $\mathfrak{P}$*-cyclic EQ-module, because for* $\varsigma \in E_\Im$ *and* $\varsigma \in \mathfrak{P}$ *we have* $\neg\varsigma\mathfrak{M}_\Im = \mu\mathfrak{M}_\Im = \{\varsigma, 1\} \subseteq E_\Im\varsigma = \{\varsigma, 1\}$

Example 7. *In Example 2, let* $\mathfrak{P} = \{0, \varsigma\}$. *Then* $\mathfrak{P}$ *is a maximal ideal of* $\Im_E$ *and* $\mathbb{Z}$ *is not a* $\mathfrak{P}$*-cyclic EQ-module, because for* $0 \in \mathbb{Z}$ *we have* $\neg 0\mathbb{Z} = 1\mathbb{Z} = \mathbb{Z} \nsubseteq E_\Im = 0$ *and* $\neg\varsigma\mathbb{Z} = \iota\mathbb{Z} = \mathbb{Z} \nsubseteq E_\Im\nu = \{\nu\}$ *and* $\neg\varsigma\mathbb{Z} = \iota\mathbb{Z} = \mathbb{Z} \nsubseteq E_\Im\nu = \{\nu\}$.

In what follows we assume that $\Im_E$ is a bounded lattice ordered EQ-algebra.
Let $\mathfrak{M}_\Im$ be an EQ-module and $\mathfrak{P}$ be a maximal ideal of $\Im_E$. Then we define

$$T_\mathfrak{P}(\mathfrak{M}_\Im) = \{\mu \in \mathfrak{M}_\Im \mid \neg\varrho\mu = 0 \text{ for some } \varrho \in \mathfrak{P}\}$$

Proposition 11. *Let* $\Im_E$ *be an IEQ-algebra,* $\mathfrak{M}_\Im$ *be an EQ-module such that* $0\mu = 0$, *for all* $\mu \in \mathfrak{M}_\Im$ *and* $\mathfrak{P}$ *be a maximal ideal of* $\Im_E$. *Then* $T_\mathfrak{P}(\mathfrak{M}_\Im)$ *is a submodule of* $\mathfrak{M}_\Im$.

Proof: Clearly $T_\mathfrak{P}(\mathfrak{M}_\Im)$ is non-empty set. Let $\mu_1, \mu_2 \in T_\mathfrak{P}(\mathfrak{M}_\Im)$. Then there exist $\varrho, \mathfrak{q} \in \mathfrak{P}$ such that $\neg\varrho\mu_1 = 0$ and $\neg\mathfrak{q}\mu_2 = 0$. We have $\varrho \boxplus_\Im \mathfrak{q} \in \mathfrak{P}$ and by Lemma 2, $\neg(\varrho \boxplus_\Im \mathfrak{q})(\mu_1 - \mu_2) = \neg(\neg\varrho \multimap_\Im \mathfrak{q})(\mu_1 - \mu_2) = (\mu_1 - \mu_2) - ((\neg\varrho \multimap_\Im q)(\mu_1 - \mu_2)) = (\mu_1 - \mu_2) - ((\neg p \wedge_\Im q \sim_\Im \neg\varrho)\mu_1 - (\neg q \wedge_\Im \varrho \sim_\Im \neg\mathfrak{q})\mu_2) = (\mu_1 - \mu_2) - ((\mu_1 - ((\neg\varrho \wedge_\Im \mathfrak{q})\mu_1 + \neg\varrho\mu_1) - (\mu_2 - (\neg\varrho \wedge_\Im \mathfrak{q})\mu_2 + \neg\mathfrak{q}\mu_2)) = 0$. Thus $\mu_1 - \mu_2 \in \mathfrak{M}_\Im$. Also, if $\zeta \in E_\Im$ and $\mu \in T_\mathfrak{P}(\mathfrak{M}_\Im)$, then there exists $\varrho \in \mathfrak{P}$ such that $\neg\varrho\mu = 0$. Hence $\neg\varrho(\zeta\mu) = \zeta(\neg\varrho\mu) = \zeta 0 = 0$, that is $\zeta\mu \in T_\mathfrak{P}(\mathfrak{M}_\Im)$. Therefore, $T_\mathfrak{P}(\mathfrak{M}_\Im)$ is a submodule of $\mathfrak{M}_\Im$. ∎

Theorem 2. *Let* $\mathfrak{M}_\Im$ *be an IEQ-module such that* $0\mu = 0$, *for all* $\mu \in \mathfrak{M}_\Im$. *Then* $\mathfrak{M}_\Im$ *is multiplication EQ-module if and only if for any maximal ideal* $\mathfrak{P}$ *of* $\Im_E$, $\mathfrak{M}_\Im = T_\mathfrak{P}(\mathfrak{M}_\Im)$ *or* $\mathfrak{M}_\Im$ *is* $\mathfrak{P}$*-cyclic.*

Proof: Let $\mathfrak{M}_\Im$ be a multiplication EQ-module and $\mathfrak{P}$ be a maximal ideal of $\Im_E$. Then, if $\mathfrak{M}_\Im = \mathfrak{P}\mathfrak{M}_\Im$ and $\mu \in \mathfrak{M}_\Im$, then there exists an ideal $\mathfrak{I}$ of $\Im_E$ such that $\mathfrak{I}\mathfrak{M}_\Im = E_\Im\mu$. Thus $E_\Im\mu = \mathfrak{I}\mathfrak{M}_\Im = \mathfrak{I}(\mathfrak{P}\mathfrak{M}_\Im) = \mathfrak{P}(\mathfrak{I}\mathfrak{M}_\Im) = \mathfrak{P}(E_\Im\mu) = \mathfrak{P}\mu$ and so there exists $\varrho \in \mathfrak{P}$ such that $\mu = \varrho\mu$. Hence $\neg\varrho\mu = (\varrho \sim_\Im 0)\mu = \mu - (\varrho\mu + 0\mu) = 0$ that is $\mu \in T_\mathfrak{P}(\mathfrak{M}_\Im)$. Therefore, $T_\mathfrak{P}(\mathfrak{M}_\Im) = \mathfrak{M}_\Im$.
Now, if $\mathfrak{P}\mathfrak{M}_\Im \subsetneq \mathfrak{M}_\Im$, then there exists $\varpi \in \mathfrak{M}_\Im$ such that $\varpi \notin \mathfrak{P}\mathfrak{M}_\Im$. Since $\mathfrak{M}_\Im$ is a multiplication EQ-module, then there exists an ideal $\mathfrak{I}$ of $\Im_E$ such that $\mathfrak{I}\mathfrak{M}_\Im = E_\Im\varpi$. If $\mathfrak{I} \subseteq \mathfrak{P}$, then $E_\Im\varpi = \mathfrak{I}\mathfrak{M}_\Im \subseteq \mathfrak{P}\mathfrak{M}_\Im$ and so $\varpi \in \mathfrak{P}\mathfrak{M}_\Im$, which is a contradiction. Hence $\mathfrak{I} \nsubseteq \mathfrak{P}$. thus there exists $\varsigma \in \mathfrak{I}$ such that $\varsigma \notin \mathfrak{P}$. since $\mathfrak{P}$ is a maximal ideal of $\Im_E$, then $\neg\varsigma \in \mathfrak{P}$. Set $\varrho = \neg\varsigma$, thus $\neg\varrho \in \mathfrak{I}$ and $(\neg\varrho)\mathfrak{M}_\Im \subseteq \mathfrak{I}\mathfrak{M}_\Im = E_\Im\varpi$. Hence $\mathfrak{M}_\Im$ is a $\mathfrak{P}$-cyclic.
Conversely, let $\mathfrak{N}_\Im$ be a submodule of $\mathfrak{M}_\Im$ and $\mathfrak{I} = (\mathfrak{N}_\Im : \mathfrak{M}_\Im)$. Then clearly $\mathfrak{I}\mathfrak{M}_\Im \subseteq \mathfrak{N}_\Im$. Suppose that, $\nu \in \mathfrak{N}_\Im$. Set $\mathfrak{K} = \{\zeta \in E_\Im \mid \zeta\nu \in \mathfrak{I}\mathfrak{M}_\Im\}$. Let $\mathfrak{K} \neq E_\Im$. Then there exists a maximal ideal $\mathfrak{Q}$ of $\Im_E$ such that $\mathfrak{K} \subseteq \mathfrak{Q}$. If $\mathfrak{M}_\Im$ is $\mathfrak{Q}$-torsion, then for each $\mu \in \mathfrak{M}_\Im$ especially $\nu \in \mathfrak{M}_\Im$ there exists $\mathfrak{s} \in \mathfrak{Q}$ such that $(\neg\mathfrak{s})\nu = 0$. Hence $\neg\mathfrak{s} \in \mathfrak{K} \subseteq \mathfrak{Q}$, which is a contradiction. Thus by assumption $\mathfrak{M}_\Im$ is $\mathfrak{Q}$-cyclic, i.e. there exists $\mathfrak{u} \in \mathfrak{M}_\Im$ and $\mathfrak{q} \in \mathfrak{Q}$ such that $(\neg\mathfrak{q})\mathfrak{M}_\Im \subseteq E_\Im\mathfrak{u}$. It follows that $(\neg\mathfrak{q})\mathfrak{N}_\Im$ is a submodule of $E_\Im\mathfrak{u}$. Set $\mathfrak{J} = \{\zeta \in E_\Im \mid \zeta\mathfrak{u} \in (\neg\mathfrak{q})\mathfrak{N}_\Im\}$. Obviously, $\mathfrak{J} \neq \emptyset$. Let $\zeta, \varpi \in \mathfrak{J}$. Then $(\zeta \boxplus_\Im \varpi)\mathfrak{u} = (\neg\zeta \multimap_\Im \varpi)\mathfrak{u} = (\neg\zeta \wedge_\Im \varpi \sim_\Im \neg\zeta)\mathfrak{u} = \mathfrak{u} - ((\neg\zeta \wedge_\Im \varpi)\mathfrak{u}) + (\neg\zeta\mathfrak{u})) = \mathfrak{u} - (\neg\zeta(\varpi\mathfrak{u}) + (\neg\zeta\mathfrak{u})) = \mathfrak{u} - (\varpi\mathfrak{u} - \zeta(\varpi\mathfrak{u}) + \mathfrak{u} - \zeta\mathfrak{u}) = -\varpi\mathfrak{u} + \zeta(\varpi\mathfrak{u}) + \zeta\mathfrak{u} = \zeta(\varpi\mathfrak{u}) + \zeta\mathfrak{u} - \varpi\mathfrak{u} \in \neg\mathfrak{q}\mathfrak{N}_\Im$. Thus $\zeta \boxplus_\Im \varpi \in \mathfrak{J}$. Let $\zeta \leq_\Im \varpi$ and $\varpi \in \mathfrak{J}$. Then $\varpi\mathfrak{u} \in \neg\mathfrak{q}\mathfrak{N}_\Im$. Hence $\zeta\mathfrak{u} = (\zeta \wedge_\Im \varpi)\mathfrak{u} = \zeta(\varpi\mathfrak{u}) \in \neg\mathfrak{q}\mathfrak{N}_\Im$, that is $\zeta \in \mathfrak{J}$. Hence $\mathfrak{J}$ is an ideal of $\Im_E$ and $(\neg\mathfrak{q})\mathfrak{N}_\Im = \mathfrak{J}\mathfrak{u}$. Also, we have $((\neg\mathfrak{q})\mathfrak{J})\mathfrak{M}_\Im = \mathfrak{J}((\neg\mathfrak{q})\mathfrak{M}_\Im) \subseteq \mathfrak{J}(E_\Im\mathfrak{u}) = \mathfrak{J}\mathfrak{u} \subseteq \mathfrak{N}_\Im$ and so $(\neg\mathfrak{q})\mathfrak{J} \subseteq \mathfrak{I}$. Hence $(\neg\mathfrak{q})\nu = ((\neg\mathfrak{q}) \wedge_\Im (\neg\mathfrak{q}))\nu = (\neg\mathfrak{q})((\neg\mathfrak{q})\nu) \in (\neg\mathfrak{q})((\neg\mathfrak{q})\mathfrak{N}_\Im) = (\neg\mathfrak{q})(\mathfrak{J}\mathfrak{u}) \subseteq \mathfrak{I}\mathfrak{M}_\Im$. So, $\neg\mathfrak{q} \in \mathfrak{K} \subseteq \mathfrak{Q}$ or $0 = \mathfrak{q} \wedge_\Im (\neg\mathfrak{q}) \in \mathfrak{Q}$, which is a contradiction. Therefore, $\mathfrak{M}_\Im$ is a multiplication module. ∎

IV. CONCLUTION

In this paper, we continued the study of EQ-modules, begun by B. Ganji Saffar (2020). The main goal of this work was to

introduce P-torsion EQ-modules and P-cyclic EQ-modules, for all maximal ideal P of $\Im_E$.

In classical ring theory there are relations between multiplication modules and P-torsion modules and P-cyclic modules. Thus we define the concepts of the maximal ideal on EQ-algebras, then we define concepts of P-torsion EQ-modules and P-cyclic EQ-modules, for each maximal ideal P of $\Im_E$ and gain interesting relation between these concepts and multiplicative EQ-modules.

REFERENCES

[1] R.A. Borzooei, J. Shohani and M. Jafari, "Extended BCK-Module," World Applied Science Journal **14-12** (2011) 1843-1850.

[2] EL-Zekey, M. Novák, V. Mestar, R. "On good EQ-algebras," Fuzzy Sets and Systems, vol. 178, pp. 1-23, 2011.

[3] B. Ganji Saffar, "EQ-modules," Math. Slovaca vol.70, No. 6, pp. 1317-1332, 2020.

[4] Hájek, P., "Fuzzy logics with noncommutative conjuctions," Journal of Logic and Computation vol. 13 pp. 469-479, 2003.

[5] Jie Qiong Shi, Xiao Long Xin, "Ideal theory on EQ-algebras", AIMS Mathematics, Vol. 6, No. 11, pp. 11686-11707, (2021).

[6] Novák, V. " EQ-algebra, Primary concepts and properties," Int. Joint, Czech Republic-Japan and Taiwan-Japan Symposium, Kitakyushu, Japan, pp. 219-223, August 2006.

[7] Novák, V. DE Baets, B. "EQ-algebras," Fuzzy Sets and Systems, vol. 160, pp. 2956-2978, 2009.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

On the Completeness of a quantale-valued metric space

Gunther Jäger

School of Mechanical Engineering, *University of Applied Sciences Stralsund*, 18435 Stralsund, Germany,
gunther.jaeger@hochschule-stralsund.de

***Abstract*— Quantale-valued metric spaces generalize both metric spaces and probabilistic metric spaces. In this paper, we study a completeness concept for quantale-valued metric spaces from the viewpoint of quantale-valued Cauchy tower spaces.**

***Index Terms*— Quantale-valued metric space, quantale-valued Cauchy tower space, completeness.**

I. Introduction

Cauchy spaces, introduced in [11], are a natural setting for studying completeness and completion [22]. Probabilistic Cauchy spaces were introduced in [20], [23] as certain "towers indexed by $[0,1]$". They were meant to be related to probabilistic metric spaces [24], however, this remained unclear in [23] and [20]. Generalizing the index set to a quantale, a quantale-valued metric space, in particular a probabilistic metric space, possesses a "natural" Cauchy tower structure [10]. We discuss the completeness of a quantale-valued metric space using this Cauchy tower structure, generalizing a result obtained in [10] about the Cauchy completeness of a quantale-valued metric space.

II. Preliminaries

For a complete lattice L with top element $\top$ and bottom element $\bot$ we call, for $\alpha, \beta \in L$, α *well-below* β, and write $\alpha \lhd \beta$, if for all subsets $D \subseteq L$ such that $\beta \leq \bigvee D$ there is $\delta \in D$ with $\alpha \leq \delta$. Then $\alpha \leq \beta$ whenever $\alpha \lhd \beta$ and $\alpha \lhd \bigvee_{\beta \in B} \beta$ if and only if $\alpha \lhd \beta$ for some $\beta \in B$. Similarly, we call α *way-below* β, and write $\alpha \prec \beta$, if for all directed subsets $D \subseteq L$ such that $\beta \leq \bigvee D$ there is $\delta \in D$ with $\alpha \leq \delta$. The way-below relation has similar properties as the well-below relation if arbitrary subsets are replaced by directed subsets. Moreover have $\alpha \vee \beta \prec \gamma$ if $\alpha, \beta \prec \gamma$, [6].

If we have $\alpha = \bigvee\{\beta \ : \ \beta \lhd \alpha\}$ for any $\alpha \in L$ then the complete lattice L is completely distributive and if we have $\alpha = \bigvee\{\beta \ : \ \beta \prec \alpha\}$ for any $\alpha \in L$, then L *continuous* [6]. For more details we refer to [6].

A *commutative and integral quantale* is a triple $\mathsf{L} = (L, \leq, *)$ with a complete lattice $(L, \leq)$ and a commutative semigroup $(L, *)$ with $\alpha * \top = \alpha$ for all $\alpha \in L$ and such that the quantale operation distributes over joins in the following sense:

$$(\bigvee_{i \in J} \alpha_i) * \beta = \bigvee_{i \in J} (\alpha_i * \beta)$$

for all $\alpha_i, \beta \in L, i \in J$. The *implication operator* in a quantale is defined by $\alpha \to \beta = \bigvee\{\gamma \in L \ : \ \alpha * \gamma \leq \beta\}$ and we have the characterization $\gamma \leq \alpha \to \beta$ if and only if $\gamma * \alpha \leq \beta$.

In this paper $\mathsf{L} = (L, \leq, *)$ is always a commutative and integral quantale with underlying completely distributive – and hence continuous – lattice.

We give three important examples:

1) The unit interval $[0,1]$ with a left-continuous *t-norm* [24].
2) The interval $[0, \infty]$ with the opposite order and addition $\alpha * \beta = \alpha + \beta$, extended by $\alpha + \infty = \infty + \alpha = \infty$, (*Lawvere's quantale*) [5].
3) A *distance distribution function* $\varphi : [0, \infty] \longrightarrow [0,1]$ satisfies $\varphi(x) = \sup_{y<x} \varphi(y)$ for all $x \in [0, \infty]$ and Δ^+ denotes the set of distance distribution functions. Ordered pointwisely, Δ^+ is a completely distributive lattice [5] and a quantale operation on Δ^+ is called a *sup-continuous triangle function* in [24].

Lemma 2.1: [10] Let $\mathsf{L} = (L, \leq, *)$ be a quantale where $(L, \leq)$ is a continuous lattice and let $\alpha \in L$. Then

1) $\bigvee_{\delta \prec \alpha}(\delta * \delta) = \alpha * \alpha$.
2) If $\epsilon \prec \alpha * \alpha$, then there is $\delta \prec \alpha$ such that $\epsilon \prec \delta * \delta$.

The power set of a set X is denoted by $\mathsf{P}(X)$ and the set of all filters on X by $\mathsf{F}(X)$. We order $\mathsf{F}(X)$ by set inclusion. For $x \in X$ then $[x] = \{A \subseteq X \ : \ x \in A\} \in \mathsf{F}(X)$. The image of a filter $\mathbb{F} \in \mathsf{F}(X)$ under a mapping $f : X \longrightarrow Y$ is defined by $f(\mathbb{F}) = \{G \subseteq Y \ : \ f(F) \subseteq G \text{ for some } F \in \mathbb{F}\} \in \mathsf{F}(Y)$.

For notions from category theory we refer to [1].

III. L-convergence tower spaces, L-Cauchy tower spaces and L-metric spaces

Definition 3.1: An L-*convergence tower space* $(X, \overline{q} = (q_\alpha)_{\alpha \in L})$ [8], is a set X with an L-*convergence tower* $(q_\alpha \colon \mathsf{F}(X) \longrightarrow \mathsf{P}(X))_{\alpha \in L}$ such that

(LC1) $x \in q_\alpha([x])$, $\forall x \in X, \alpha \in L$;
(LC2) $\forall \mathbb{F}, \mathbb{G} \in \mathsf{F}(X), \alpha \in L$, $q_\alpha(\mathbb{F}) \subseteq q_\alpha(\mathbb{G})$ whenever $\mathbb{F} \leq \mathbb{G}$;
(LC3) $\forall \alpha, \beta \in L, \mathbb{F} \in \mathsf{F}(X)$, $q_\beta(\mathbb{F}) \subseteq q_\alpha(\mathbb{F})$ whenever $\alpha \leq \beta$;
(LC4) $x \in q_\bot(\mathbb{F})$, $\forall x \in X$, $\mathbb{F} \in \mathsf{F}(X)$.

If $(X, \overline{q})$ satisfies $x \in q_{\vee A}(\mathbb{F})$ whenever $x \in q_\alpha(\mathbb{F}) \ \ \forall \alpha \in A$, it is called *left-continuous*. A mapping $f \colon (X, \overline{q}) \longrightarrow (X', \overline{q'})$ between L-convergence tower spaces is called *continuous* if, for all $x \in X$, and for all $\mathbb{F} \in \mathsf{F}(X)$, we have $f(x) \in q'_\alpha(f(\mathbb{F}))$ whenever $x \in q_\alpha(\mathbb{F})$. The category of L-convergence tower spaces and continuous mappings is denoted by L-CTS.

Examples are classical convergence spaces [3] for $L = \{0, 1\}$; limit tower spaces [2] – or in the left-continuous case approach limit spaces [15] – for Lawvere's quantale $\mathsf{L} = ([0, \infty], \geq, +)$; probabilistic convergence spaces [23] for $\mathsf{L} = ([0, 1], \leq, *)$ and [7] for $\mathsf{L} = (\Delta^+, \leq, *)$.

Definition 3.2 ([10]): An L-*Cauchy tower space* $(X, \overline{C}) = (X, (C_\alpha)_{\alpha \in L})$ is set X and an L-*Cauchy tower*, i.e. $C_\alpha \subseteq \mathsf{F}(X)$ for all $\alpha \in L$, such that

(LChy1) $[x] \in C_\alpha$ for all $x \in X, \alpha \in L$;
(LChy2) $\mathbb{G} \in C_\alpha$ whenever $\mathbb{G} \geq \mathbb{F} \in C_\alpha$;
(LChy3) $\mathbb{F} \in C_\alpha$ whenever $\alpha \leq \beta, \mathbb{F} \in C_\beta$;
(LChy4) $C_\bot = \mathbb{F}(X)$.
(LChy5) $\mathbb{F} \wedge \mathbb{G} \in C_{\alpha * \beta}$ whenever $\mathbb{F} \in C_\alpha, \mathbb{G} \in C_\beta, \mathbb{F} \vee \mathbb{G}$ exists.

If $(X, \overline{C})$ satisfies $\mathbb{F} \in C_{\bigvee A}$ whenever $\mathbb{F} \in C_\alpha$ for all $\alpha \in A \subseteq L$, it is called *left-continuous*. A mapping $f : (X, \overline{C}) \longrightarrow (X', \overline{C'})$ is called *Cauchy-continuous* if $f(\mathbb{F}) \in C'_\alpha$ whenever $\mathbb{F} \in C_\alpha$. The category of L-Cauchy tower spaces and Cauchy-continuous mappings is denoted by L-ChyTS.

Examples are probabilistic Cauchy spaces [19] for $\mathsf{L} = ([0, 1], \leq, *)$ with a continous t-norm $*$ and [12] for $* = \min$; Cauchy tower spaces [18] – and in the left-continuous case approach Cauchy spaces [17] – for Lawever's quantale.

As $\top * \top = \top$, for an L-Cauchy tower space $(X, \overline{C})$, $(X, C_\top)$ is Cauchy space in the classical definition [11].

An L-Cauchy tower space $(X, \overline{C})$ has an underlying L-convergence tower space $(X, \overline{q^{\overline{C}}})$ defined by $x \in q^{\overline{C}}_\alpha(\mathbb{F})$ if and only if $\mathbb{F} \wedge [x] \in C_\alpha$, see [10].

Definition 3.3: An L-*metric space* [5], [14] is a pair (X, d) of a set X and L-*metric* $d : X \times X \longrightarrow L$ such that

(LM1) $d(x, x) = \top \ \forall x \in X$;
(LM2) $d(x, y) * d(y, z) \leq d(x, z) \ \forall x, y, z \in X$.

It is called *symmetric* if moreover

(LMs) $d(x, y) = d(y, x) \ \forall x, y \in X$.

A mapping between two L-metric spaces, $f : (X, d) \longrightarrow (X', d')$ is called an L-*metric morphism* if $d(x_1, x_2) \leq d'(f(x_1), f(x_2))$ for all $x_1, x_2 \in X$. The category of L-metric spaces and L-metric morphisms is denoted by L-MET, the subcategory of symmetric L-metric spaces is denoted by L-sMET.

Examples are preordered sets for $\mathsf{L} = (\{0, 1\}, \leq, \wedge)$; quasi-metric spaces for Lawvere's quantale and probabilistic quasi-metric spaces for $\mathsf{L} = (\Delta^+, \leq, *)$. Fuzzy metric spaces in the definition of Kramosil and Michálek [13] can be identified with probabilistic metric spaces, see e.g. [9].

For an L-metric space (X, d), the L-convergence tower $\overline{q^d}$ is defined by [8]

$$x \in q^d_\alpha(\mathbb{F}) \iff \bigvee_{F \in \mathbb{F}} \bigwedge_{y \in F} d(x, y) \geq \alpha.$$

In a similary way, an L-Cauchy tower, $\overline{C^d}$, is defined by [10]

$$\mathbb{F} \in C^d_\alpha \iff \bigvee_{F \in \mathbb{F}} \bigwedge_{x, y \in F} d(x, y) \geq \alpha.$$

Proposition 3.4 ([10]): Let $(X, d) \in |\mathsf{L}\text{-MET}|$. Then $q^{\overline{C^d}}_\alpha(\mathbb{F}) \subseteq q^d_\alpha(\mathbb{F})$ and if (X, d) is symmetric, we have $q^d_\alpha(\mathbb{F}) \subseteq q^{\overline{C^d}}_{\alpha * \alpha}(\mathbb{F})$.

For an L-MET-morphism $f : (X, d) \longrightarrow (X', d')$ we obtain an L-ChyTS-morphism $f : (X, \overline{C^d}) \longrightarrow (X', \overline{C^{d'}})$, and this defines a functor $\mathsf{F} : \mathsf{L}\text{-MET} \longrightarrow \mathsf{L}\text{-ChyTS}$.

For $(X, \overline{C}) \in |\mathsf{L}\text{-ChyTS}|$ we define $d^{\overline{C}} : X \times X \longrightarrow L$ by

$$d^{\overline{C}}(x, y) = \bigvee_{[x] \wedge [y] \in C_\alpha} \alpha$$

for all $x, y \in X$, [10]. Then $(X, d^{\overline{C}}) \in |\mathsf{L}\text{-sMET}|$ and we have $d^{\overline{C^d}} = d$. For a Cauchy continuous mapping $f : (X, \overline{C}) \longrightarrow (X', \overline{C'})$ then $f : (X, d^{\overline{C}}) \longrightarrow (X', d^{\overline{C'}})$ is an L-MET-morphism. Hence we have a functor $\mathsf{G} : \mathsf{L}\text{-ChyTS} \longrightarrow \mathsf{L}\text{-sMET}$. The functor G is left-inverse for F, i.e. we have $\mathsf{G} \circ \mathsf{F} = id_{\mathsf{L}\text{-sMET}}$. We note that L-sMET is in general not a reflective subcategory of L-ChyTS ([16], Remark 7.9).

IV. A COMPLETENESS NOTION FOR L-METRIC SPACES

In this section we extend a result obtained in [10], making use of the full L-Cauchy tower of an L-metric space, while in [10] we used only the "top level" $C^d_\top$.

We call an L-metric space (X,d) *strongly Cauchy-complete* if for all $\alpha \in L$, for $\mathbb{F} \in C^d_\alpha$ there is $a \in X$ such that $a \in q^{d^s}_{\alpha * \alpha}(\mathbb{F})$. It is called *Cauchy complete* if this holds true only for $\alpha = \top$, i.e. if $\mathbb{F} \in C^d_\top$ implies that there is $a \in X$ such that $a \in q^{d^s}_\top(\mathbb{F})$. Here $d^s : X \times X \longrightarrow L$, defined by $d^s(x,y) = d(x,y) * d(y,x)$ for all $x, y \in X$, is the *symmetrization* of d.

We review concepts introduced in [4]. Let $(X,d) \in |\mathsf{L\text{-}MET}|$. A mapping $\Phi : X \longrightarrow L$ is an *order ideal* if $d(y,x) * \Phi(x) \leq \Phi(y)$ for all $x, y \in X$ and it is an *order filter* if $\Phi(x) * d(x,y) \leq \Phi(y)$ for all $x, y \in X$. Using the implication operator we have for an order ideal $d(y,x) \leq \Phi(x) \to \Phi(y)$ and $d(x,y) \leq \Phi(x) \to \Phi(y)$ for an order filter. If we define the L-metric $d_{\mathsf{L}} : L \times L \longrightarrow L$ by $d_{\mathsf{L}}(\alpha, \beta) = \alpha \to \beta$ for $\alpha, \beta \in L$, an order ideal becomes an L-metric morphism from (X, d^{op}) to (L, d_{L}) and an L-order filter becomes an L-metric morphism from (X,d) to (L, d_{L}).

Lemma 4.1: Let (X,d) be an L-metric space.

1) For $a \in X$, $\downarrow(a)$ defined by $\downarrow(a)(x) = d(x,a)$ for all $x \in X$ is an order ideal and $\uparrow(a)$ defined by $\uparrow(a)(x) = d(a,x)$ for all $x \in X$ is an order filter [4].
2) For a filter $\mathbb{F}$ on X, $\Phi : X \longrightarrow L$ defined by $\Phi(x) = \bigvee_{F \in \mathbb{F}} \bigwedge_{y \in F} d(x,y)$ for all $x \in X$ is an order ideal and $\Psi : X \longrightarrow L$ defined by $\Psi(x) = \bigvee_{F \in \mathbb{F}} \bigwedge_{y \in F} d(y,x)$ for all $x \in X$ is an order filter [10].

We leave the simple proof of the next lemma to the reader.

Lemma 4.2: Let (X,d) be an L-metric space and let $\Phi : X \longrightarrow L$ be an order ideal and $\Psi : X \longrightarrow L$ be an order filter. For $\alpha \in L$ we define $\alpha\Phi : X \longrightarrow L$ by $\alpha\Phi(x) = \alpha * \Phi(x)$ for all $x \in X$ and likewise $\alpha\Psi : X \longrightarrow L$ by $\alpha\Psi(x) = \alpha * \Psi(x)$ for all $x \in X$. Then $\alpha\Phi$ is an order ideal and $\alpha\Psi$ is an order filter.

Definition 4.3 ([4] for $\alpha = \top$): Let (X,d) be an L-metric space and let $\alpha \in L$. A pair (Φ, Ψ) with an order ideal $\Phi : X \longrightarrow L$ and an order filter $\Psi : X \longrightarrow L$ is called an *α-cut* on X if $\alpha * \alpha \leq \bigvee_{x \in X} \Phi(x) * \Psi(x)$ and $\Phi(x) * \Psi(y) \leq d(x,y)$ for all $x, y \in X$.

If $\alpha = \top$, we simply speak of a *cut* [4].

It is not difficult to see that for each $a \in X$, the pair $(\alpha \downarrow (a), \alpha \uparrow (a))$ is an α-cut on X.

For an α-cut (Φ, Ψ) we say that a point $a \in X$ *represents the α-cut* if $\alpha\Phi = \alpha \downarrow (a)$ and $\alpha\Psi = \alpha \uparrow (a)$, i.e. if $\alpha * \Phi(x) = \alpha * d(x,a)$ as well as $\alpha * \Psi(x) = \alpha * d(a,x)$ for all $x \in X$.

Proposition 4.4: Let (X,d) be an L-metric space. If for all $\alpha \in L$ and for all α-cuts (Φ, Ψ) there is $a \in X$ that represents the α-cut, then (X,d) is strongly Cauchy complete.

Proof: Let $\mathbb{F} \in C^d_\alpha$ and consider the order ideal Φ and the order filter Ψ defined in Lemma 4.1(2). Let $\epsilon \prec \alpha * \alpha$. By virtue of Lemma 2.1, we can choose $\delta \prec \alpha$ such that $\epsilon \prec \delta * \delta$. Then $\delta \prec \alpha \leq \bigvee_{F \in \mathbb{F}} \bigwedge_{x,y \in F} d(x,y)$. Hence there is $F \in \mathbb{F}$ such that $d(x,y) \succ \delta$ as well as $d(y,x) \succ \delta$ for all $x, y \in F$. We conclude for all $x \in F$

$$\begin{aligned} \epsilon \prec \delta * \delta \ &\leq \ \bigwedge_{y \in F} d(x,y) * \bigwedge_{y \in F} d(y,x) \\ &\leq \bigvee_{F \in \mathbb{F}} \bigwedge_{y \in F} d(x,y) * \bigvee_{F \in \mathbb{F}} \bigwedge_{y \in F} d(y,x) \ = \ \Phi(x) * \Psi(x). \end{aligned}$$

Hence $\epsilon \prec \bigvee_{x \in X} \Phi(x) * \Psi(x)$ and as $\epsilon \prec \alpha * \alpha$ was arbitrary we obtain $\alpha * \alpha \leq \bigvee_{x \in X} \Phi(x) * \Psi(x)$. Furthermore, we have

$$\begin{aligned} \Phi(x) * \Psi(y) &= \bigvee_{F \in \mathbb{F}} \bigwedge_{z \in F} d(x,z) * \bigvee_{G \in \mathbb{F}} \bigwedge_{z \in G} d(z,y) \\ &\leq \bigvee_{F,G \in \mathbb{F}} \bigwedge_{z \in F \cap G} d(x,z) * d(z,y) \\ &\leq \ d(x,y). \end{aligned}$$

Therefore (Φ, Ψ) is an α-cut on X. By the assumption, there is $a \in X$ which represents the cut. Then $\alpha = \alpha * \Phi(a) \leq \bigvee_{F \in \mathbb{F}} \bigwedge_{z \in F} d(a,z)$ and $\alpha = \alpha * \Psi(a) \leq \bigvee_{F \in \mathbb{F}} \bigwedge_{z \in F} d(z,a)$. Let $\epsilon \prec \alpha$. There is a set $F \in \mathbb{F}$ such that for all $z \in F$ we have $d(a,z) \geq \epsilon$ and there is a set $G \in \mathbb{F}$ such that for all $z \in G$ we have $d(z,a) \geq \epsilon$. We conclude $\epsilon * \epsilon \leq d^s(z,a)$ for all $z \in F \cap G$ and thus $\epsilon * \epsilon \leq \bigvee_{H \in \mathbb{F}} \bigwedge_{z \in H} d^s(z,a)$. Again, as $\epsilon \prec \alpha$ was arbitrary, we obtain $\alpha * \alpha \leq \bigvee_{H \in \mathbb{F}} \bigwedge_{z \in H} d^s(z,a)$ and therefore $a \in q^{d^s}_{\alpha * \alpha}(\mathbb{F})$. □

A partial converse is the following result. We need that in the lattice $(L, \leq)$ the way-below relation and the well-below relation coincide. This is e.g. the case for a linearly ordered lattice.

Proposition 4.5: Let in $\mathsf{L} = (L, \leq, *)$ the the way-below relation and the well-below relation coincide and let (X,d) be a strongly complete L-metric space. Then, for each idempotent $\alpha \in L$, and each α-cut (Φ, Ψ) there is a point $a \in X$ that represents the α-cut.

Proof: Let $\alpha \in L$ be idempotent, i.e. we have $\alpha = \alpha * \alpha$. We consider an α-cut (Φ, Ψ). For $\epsilon \prec \alpha$ we define the set $F_\epsilon = \{x \in : \ \Phi(x) * \Psi(x) \succ \epsilon\}$. From the idempotency of α we get $\alpha = \alpha * \alpha \leq \bigvee_{x \in X} \Phi(x) * \Psi(x)$ and we conclude, using $\varepsilon \prec \alpha \iff \varepsilon \lhd \alpha$, that F_ϵ is not empty. Therefore $\{F_\epsilon : \epsilon \prec \alpha\}$ is a filter basis. We denote the generated filter by $\mathbb{F}$ and show that this filter is in C^d_α. Let $\epsilon \prec \alpha = \alpha * \alpha$.

Again by Lemma 2.1 we find $\delta \prec \alpha$ with $\epsilon \prec \delta * \delta$. Then we have for $x, y \in F_\delta$

$$\begin{aligned}\delta * \delta &= \Phi(x) * \Psi(x) * \Phi(y) * \Psi(y)\\ &\leq \Phi(x) * \Psi(y)\\ &\leq d(x,y).\end{aligned}$$

We conclude $\epsilon \prec \delta * \delta \leq \bigvee_{F \in \mathbb{F}} \bigwedge_{x,y \in F} d(x,y)$. As $\epsilon \prec \alpha$ was arbitrary, we obtain $\mathbb{F} \in C^d_\alpha$.

As (X,d) is strongly Cauchy complete and α is idempotent, there is $a \in X$ with $a \in q^{d^s}_{\alpha * \alpha}(\mathbb{F}) = q^{d^s}_\alpha(\mathbb{F})$. This implies

$$\alpha \leq \bigvee_{\epsilon \prec \alpha} \bigwedge_{z \in F_\epsilon} d(a,z) * d(z,a) \leq \bigvee_{\epsilon \prec \alpha} \bigwedge_{z \in F_\epsilon} d(z,a)$$

and as $\Psi : (X,d) \longrightarrow (L, d_L)$ is an L-MET-morphism we conclude

$$\alpha \leq \bigvee_{\epsilon \prec \alpha} \bigwedge_{z \in F_\epsilon} (\Psi(z) \to \Psi(a)) = \bigvee_{\epsilon \prec \alpha} ((\bigvee_{z \in F_\epsilon} \Psi(z)) \to \Psi(a)).$$

Consider $\delta \prec \alpha$. Then there is $\epsilon_\delta \prec \alpha$ with $\delta * \bigvee_{z \in F_{\epsilon_\delta}} \Psi(z) \leq \Psi(a)$. As $\epsilon_\delta \leq \epsilon_\delta \vee \delta \prec \alpha$ we have $F_{\epsilon_\delta \vee \delta} \subseteq F_{\epsilon_\delta}$ and therefore $\delta * \bigvee_{z \in F_{\epsilon_\delta \vee \delta}} \Psi(z) \leq \Psi(a)$.

For $z \in F_{\epsilon_\delta \vee \delta}$ we have $\Psi(z) \geq \Phi(z) * \Psi(z) \geq \epsilon_\delta \vee \delta \geq \delta$ and therefore $\delta * \delta \leq \Psi(a)$. Again, as $\delta \prec \alpha$ was arbitrary, we conclude $\alpha = \alpha * \alpha \leq \Psi(a)$. In a similar way we see that $\alpha \leq \Phi(a)$. As Φ is an order ideal we conclude $\alpha * d(x,a) \leq d(x,a) * \Phi(a) \leq \Phi(x)$ for all $x \in X$. From $\Phi(x) * \alpha = \Phi(x) * \Psi(a) \leq d(x,a)$ we moreover obtain $\alpha * \Phi(x) = \alpha * \alpha * \Phi(x) \leq \alpha * d(x,a)$ for all $x \in X$. Therefore $\alpha\Phi = \alpha \downarrow (a)$. In a similar way we see that $\alpha\Psi = \alpha \uparrow (a)$ and so $a \in X$ represents the α-cut (Φ, Ψ). □

Restricting the proofs of Propositions 4.4 and 4.5 to the idempotent element $\top \in L$, we conclude that an L-metric space (X,d) is Cauchy complete if and only if each cut is represented by an element $a \in X$. In this case, in Proposition 4.5 we do not need to insist that the way-below relation equals the well-below relation but the requirement of a value quantale [4] is sufficient. This was shown already in [4].

REFERENCES

[1] J. Adámek, H. Herrlich and G.E. Strecker, "Abstract and Concrete Categories," Wiley, New York, 1989.

[2] P. Brock and D.C. Kent, "Approach spaces, limit tower spaces, and probabilistic convergence spaces," Appl. Cat. Structures 5 (1997), pp. 99 – 110.

[3] H.R. Fischer, "Limesräume," Math. Ann. 137 (1959), pp. 269 – 303.

[4] R.C. Flagg, "Completeness in continuity spaces," in: Category Theory 1991, CMS Conf. Proc. 13, pp. 183 – 199, 1992.

[5] R.C. Flagg, "Quantales and continuity spaces," Algebra Univers. 37 (1997), pp. 257 – 276.

[6] G. Gierz, K.H. Hofmann, K. Keimel, J.D. Lawson, M.W. Mislove and D.S. Scott, "Continuous lattices and domains," Cambridge University Press, 2003.

[7] G. Jäger, "A convergence theory for probabilistic metric spaces," Quaest. Math. 38 (2015), pp. 587 – 599.

[8] G. Jäger and T.M.G. Ahsanullah, "Characterization of quantale-valued metric spaces and quantale-valued partial metric spaces by convergence," Applied Gen. Topology 19, no. 1 (2018), pp. 129 – 144.

[9] G. Jäger, "The Wijsman structure of a quantale-valued metric space," Iranian J. Fuzzy Systems 17, no. 1, (2020), pp. 171 – 184.

[10] G. Jäger and T.M.G. Ahsanullah, "Quantale-valued Cauchy tower spaces and completeness," Applied Gen. Topology, to appear.

[11] H.H. Keller, "Die Limes-Uniformisierbarkeit der Limesräume," Math. Ann. 176 (1968), pp. 334 – 341.

[12] D.C. Kent and G.D. Richardson, "Completions of probabilistic Cauchy spaces," Math. Japonica 48, no. 3 (1998), pp. 399 – 407.

[13] I. Kramosil and J. Michálek, "Fuzzy metrics and statistical metric spaces," Kybernetika 11 (1975), pp. 336 – 344.

[14] F.W. Lawvere, "Metric spaces, generalized logic, and closed categories," Rendiconti del Seminario Matematico e Fisico di Milano 43 (1973), pp. 135 – 166. Reprinted in: "Reprints in Theory and Applications of Categories" 1 (2002), pp. 1 – 37.

[15] R. Lowen, "Index Analysis," Springer, London, Heidelberg, New York, Dordrecht 2015.

[16] R. Lowen and Y.J. Lee, "Approach theory in merotopic, Cauchy and convergence spaces. I," Acta Math.Hungarica 83, no. 3 (1999), pp. 189 – 207.

[17] R. Lowen and Y.J. Lee, "Approach theory in merotopic, Cauchy and convergence spaces. II," Acta Math.Hungarica 83, no. 3 (1999), pp. 209 – 229.

[18] J. Minkler, G. Minkler and G. Richardson, "Subcategories of Filter Tower Spaces," Appl. Categ. Structures 9 (2001), pp. 369 – 379.

[19] H. Nusser, "A generalization of probabilistic uniform spaces," Appl. Categ. Structures 10 (2002), pp. 81 – 98.

[20] H. Nusser, "Completion of probabilistic uniform limit spaces," Quaest. Math. 26 (2003), pp. 125 – 140.

[21] Q. Pu and D. Zhang, "Preordered sets valued in a GL-monoid," Fuzzy Sets and Systems 187 (2012), pp. 1 – 32.

[22] E.E. Reed, "Completion of uniform convergence spaces," Math. Ann. 194 (1971), pp. 83 – 108.

[23] G.D. Richardson and D.C. Kent, "Probabilistic convergence spaces," J. Austral. Math. Soc. (Series A) 61 (1996), pp. 400 – 420.

[24] B. Schweizer and A. Sklar, "Probabilistic Metric Spaces," North Holland, New York, 1983.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Online streaming feature selection based on Sugeno fuzzy integral

Amin Hashemi[1], Mohammad-Reza Pajoohan[2], Mohammad Bagher Dowlatshahi[3]

[1] *Department of Computer Engineering, Faculty of Engineering, Yazd University, Yazd, Iran*, hashemi.am@stu.yazd.ac.ir
[2] *Department of Computer Engineering, Faculty of Engineering, Yazd University, Yazd, Iran*, pajoohan@yazd.ac.ir
[3] *Department of Computer Engineering, Faculty of Engineering, Lorestan University, Khorramabad, Iran*, dowlatshahi.mb@lu.ac.ir

***Abstract*— Feature selection is a step in which an optimal subset of fe atures is selected for the learning process. This step is applied to high-dimensional data, so that irrelevant and redundant features are removed from the data. Traditional feature selection methods require the entire feature space. At the same time, in many real-world applications such as online social networks, it is impossible to acquire or wait to get the whole feature space. Therefore, online feature selection methods are considered to handle this challenge. In this paper, we present an online feature selection method based on the concept of Sugeno fuzzy integral. According to this method, the streaming features are examined based on several measures, and these measures are combined based on the Sugeno operator. If the aggregated value reaches a pre-defined threshold, the desired feature is selected; otherwise, it is not considered. To prove the performance of the proposed method, comparisons have been made with five online feature selection methods based on two classifiers. We show that the proposed method outperforms competitive methods based on the experimental results.**

***Index Terms* — Online feature selection; Sugeno fuzzy integral; streaming features; high-dimensional data.**

I. INTRODUCTION

We can call the new era the age of big data because data is produced in vast numbers and dimensions. Such large amounts of data face many problems for learning algorithms. Obtaining useful information from this data is a significant challenge in machine learning and data mining that many researchers are looking for the ways to solve these problems [1]–[5]. This issue, known as the curse of dimensionality, reduces the performance of learning models and classification accuracy, increases computational complexity and learning time, and causes over-fitting in learning models [6]–[8]. Dimensionality reduction methods are used to handle this challenge. Feature selection is one of the most effective techniques in this field [9]–[11].

High Dimensional means that the number of dimensions is staggeringly high. Some features are not relevant to the class label, and some are redundant among the data features. Thus, removing these features and selecting a small yet effective subset help to solve mentioned problems. This process is known as feature selection [9], [12]–[14].

So far, many efforts have been made to provide effective feature selection methods. However, common feature selection methods require the entire feature space. At the same time, we do not always have the entire feature space in real-world applications, and in some cases, the number of features is unlimited, and they are added continuously. For example, in the case of hot topic detection in online social networks such as Twitter and Facebook, the feature is added continuously. In this case, keywords are used as features to identify hot topics. Over time, due to changes in different situations, other hot topics arise that cause adding new keywords (features) to the feature set [15]–[17]. Every time new features arrive, we should decide what features should remain.

Therefore, online streaming feature selection (OSFS) methods emerge to select the best features in applications where data is streamed into the feature space. These methods can also be used in datasets with a specific number of attributes where data sets have a very high number of features that do not fit in the memory. In this type of feature selection, it is assumed that the number of instances is fixed, and the number of features increases over time. A feature stream is evaluated while it enters the feature space. In general, online feature selection methods based on feature streams are divided into two general categories: single-stream methods and group stream methods.

Single-stream methods are used in problems where one feature is added at a time, and in group stream methods, the feature stream is included by more than one feature. The primary approach in online feature selection methods is to decide based on a criterion to select or withdraw a feature [18].

In this paper, we present a feature selection method based on a single-stream procedure using the concept of the fuzzy integral. Accordingly, the Sugeno operator, one of the fuzzy integration methods, is considered as our criterion for selecting features in the online space. In fact, in this article, our goal is to use the combination of several criteria for feature evaluation instead of the usual online streaming feature selection methods, as one believes the variety of different criteria can lead to a better result than one criterion. Finally, we use the Sugeno fuzzy integral operator to combine the considered criteria in this approach. The value obtained from this combination is considered the feature selection measure. In this case, if this value is above a pre-defined threshold, the attribute is selected and otherwise ignored.

Five ESFS algorithms are provided to compare with the proposed method based on real-world datasets to prove the performance of the proposed method. Experiments show that the proposed method outperforms competitive methods.

The structure of this paper is organized as follows: In Section II, related methods are reviewed. Section III presents the fundamental concepts, and Section IV describes the proposed method in details. Section V presents the experimental results, and in Section VI, the conclusion is provided.

II. RELATED WORKS

This section reviews several OSFS methods that belong to the single-stream category. SOALA [19] (a scalable and accurate online feature selection approach) is an OSFS method that selects features based on pairwise comparisons. OFSMI [1] is an OSFS algorithm that evaluates the streaming features using mutual information. In this method, when each attribute enters the attribute space, its mutual information is first calculated and temporarily stored if the obtained value is more than zero. The features that pass this step will enter in redundancy level. OSFS-FI [20] is a novel OSFS algorithm based on the interaction between features. In the alpha-investigating [21] method, a p-value is computed for each streaming feature, which refers to a probability of acceptance of features. If the p-value is more significant than alpha, the feature is accepted; otherwise is ignored. K-OSFD [22] is an OSFS method based on the concept of k-nearest neighbors. In this method, the information is captured using the neighborhood rough set theory, and then the features with more separation power are chosen. OFS-3AM [23] is a non-parametric OSFS algorithm that finds the most relevant and less redundant features using an adapted neighborhood rough set. OFS-Density [24] also used rough set theory for the OSFS problem. In OFS-Density, the increase of feature dependency is the measure for selecting the features.

III. FUNDAMENTAL CONCEPTS

A. Sugeno Fuzzy Integral

In aggregation problems where we are looking to combine the values of a variable based on several criteria, fuzzy integrals can be used. In fuzzy integrals, it is assumed that the input values and the fuzzy integral are all available in a particular range. In fact, fuzzy integrals operate based on fuzzy criteria.

Definition 1. Assume that $\mathcal{N} = \{1,2, \dots, n\}$. A fuzzy measure can be defined as a set function $\mu: 2^{\mathcal{N}} \longrightarrow [0,1]$ that satisfies the following conditions [25]:

1) *Monotonic:* $If\ \mathcal{A} \subset B, then\ \mu(\mathcal{A}) \leq \mu(\mathcal{B})$
2) *Normalized:* $\mu(\emptyset) = 0\ and\ \mu(\mathcal{N}) = 1$

Definition 2. If we assume $X = \{x_1, x_2, x_3, \dots, x_n\}$ as a finite set and $\lambda \in (-1, \infty)$, then λ-fuzzy measure can be defined as a function $g_\lambda: 2^x$ that satisfies the following conditions:

- $g_\lambda(X) = 1$
- $If\ \mathcal{A}, \mathcal{B} \in 2^x, then\ g_\lambda(\mathcal{A} \cup \mathcal{B}) = g_\lambda(\mathcal{A}) + g_\lambda(\mathcal{B}) + \lambda g_\lambda(\mathcal{A}) g_\lambda(\mathcal{B})\ where\ \mathcal{A} \cap \mathcal{B} = \emptyset$

In this situation, we should calculate λ to obtain the g_λ value. The fuzzy density measure $g(X) = g_\lambda(\{x_1, x_2, x_3, \dots, x_n\})$ is calculated as follows:

$$g_\lambda(\{x_1, x_2, x_3, \dots, x_n\}) = \sum_{i=1}^{n} g_i + \lambda \sum_{i_1=1}^{n-1} \sum_{i_2=i_1+1}^{n} g_{i_1} . g_{i_2} + \cdots + \lambda^{n+1} g_1 . g_2 \dots . g_n = \frac{1}{\lambda} [\prod_{i=1}^{n} (1 + \lambda g_i) - 1], \quad (1)$$

where $\lambda \in (-1, \infty)$. As we mentioned before, $g_\lambda(X) = 1$. Thus we can simplify the above formula to calculate λ as follows:

$$\lambda = \prod_{i=1}^{n} (\lambda g_i + 1) - 1. \quad (2)$$

The proof of this formula is presented in reference [26].

Definition 3. Let μ be a fuzzy measure on X, the Sugeno integral of a function $f: X \longrightarrow [0,1]$ concerning μ is given below:

$$\int f(x) d\,\mu = \max_{1 \leq i \leq n} \left(min\left(f(x_i), \mu(\mathcal{A}_i) \right) \right), \quad (3)$$

where $f(x_i) = \{ f(x_1), f(x_2), \dots, f(x_n) \}$ indicates the ranges [27].

To better understand this method, we use a numerical example. Assume that $X = \{x_1, x_2, x_3\}$ with ranges $F(x_1) = 0.4$, $F(x_2) = 0.6$, and $F(x_3) = 0.8$. Also, the fuzzy density values are presented as $g_\lambda(\{x_1\}) = 0.4$, $g_\lambda(\{x_2\}) = 0.3$, and $g_\lambda(\{x_3\}) = 0.4$. Now we want to use Sugeno fuzzy integral to aggregate the variables value considering fuzzy density values. Thus, we need to obtain λ first. Based on Eq. 2, λ is computed as follows:

$$\lambda = (0.4\lambda + 1)(0.3\lambda + 1)(0.2\lambda + 1) - 1$$

We solve the equation and achieve $\lambda = [0, -11.87, 0.3719]$. We know that $\lambda \in (-1, \infty)$, thus the accepted values for λ are 0 and 0.3719. We solve the aggregation for $\lambda = 0.3719$. The values of g_λ are computed for all cases of X based on definition 2. These values are calculated as follows:

$$g_\lambda(\{x_1, x_2\}) = g_\lambda(\{x_1\}) + g_\lambda(\{x_2\}) + \lambda g_\lambda(\{x_1\}) g_\lambda(\{x_2\}) = 0.7446$$

$$g_\lambda(\{x_2, x_3\}) = g_\lambda(\{x_2\}) + g_\lambda(\{x_3\}) + \lambda g_\lambda(\{x_2\}) g_\lambda(\{x_3\}) = 0.5223$$

$$g_\lambda(\{x_1,x_3\}) = g_\lambda(\{x_1\}) + g_\lambda(\{x_3\}) + \lambda g_\lambda(\{x_1\})g_\lambda(\{x_3\}) = 0.7323$$

$$g_\lambda(\{x_1,x_2,x_3\}) = 1$$

The Sugeno fuzzy integral is computed using Eq.3.

$$\int f(x)d\,\mu = max\begin{pmatrix} min\big(f(x_1), g_\lambda(\{x_1,x_2,x_3\})\big), \\ min\big(f(x_2), g_\lambda(\{x_2,x_3\})\big), min\big(f(x_3), g_\lambda(\{x_3\})\big)\end{pmatrix} =$$

$$max\begin{pmatrix} min(0.4,1), \\ min(0.6,0.5223), min(0.8,0.4)\end{pmatrix} =$$

$$\max(0.4,0.5223,0.4) = 0.5223$$

B. Symmetrical Uncertainty (SU)

Symmetric uncertainty is a method of calculating how relevant an attribute is to a class label. This method is defined based on the concept of information gain and information entropy. If we consider X as a discrete variable, then the entropy of X is calculated as follows:

$$H(X) = -\sum_{x\in X} P(x)\log_2\big(P(x)\big), \quad (4)$$

where $x \in X$ and is the probability of x for all potential values in X. On the other hand, the conditional entropy between two discrete random variables X and Y is calculated as follows:

$$H(X|Y) = -\sum_{x\in X} P(y)\sum_{x\in X} P(x|y)\log_2\big(P(x|y)\big). \quad (5)$$

According to Eq. 3 and Eq. 4, Information Gain (IG) is computed:

$$IG(X|Y) = H(X) - H(X|Y), \quad (6)$$

Information gain [28] measures the relevancy between two random variables. If the value is zero for two variables, they are independent, and their dependence grows with increasing its value. The disadvantage of information gain is that it biases towards variables with diverse values. Therefore, the values are normalized in [0,1] interval to eliminate this weakness. This normal value calls Symmetrical Uncertainty (SU), and the following equation computes it:

$$SU(X,Y) = 2\left[\frac{IG(X,Y)}{H(X)+H(Y)}\right], \quad (7)$$

where 0 means the independence of two variables and 1 means their complete dependency [29].

C. Pearson Correlation Coefficient

The Pearson correlation coefficient between two variables is obtained by normalizing the covariance between them by dividing the covariance value by the product of their variance. This metric uses to capture the relevancy degree between two random variables in [-1,1] interval. The following equation is used to calculate the correlation coefficient between two variables X and Y [30]:

$$r(X,Y) = \frac{Cov(X,Y)}{S_X.S_Y}. \quad (8)$$

IV. PROPOSED METHOD

In this section, we introduce our proposed algorithm in details. This algorithm is proposed for the online streaming feature selection problem that one feature is added to the feature space at a time. We used the ensemble of three feature evaluation metrics in this method. We did the aggregation process by the Sugeno fuzzy integral operator to achieve our decision measure in online selection. Algorithm 1 shows the proposed method step-by-step.

Algorithm 1: Online streaming feature selection based on Sugeno fuzzy integral (OSFSSU)

Input: $N \times 1$ label vector $\boldsymbol{Y}$, $N \times 1$ stream feature vector $\boldsymbol{f_i}$ at time i
Output: Vector S represents the selected features
1. $\boldsymbol{S} = \emptyset$;
2. $\boldsymbol{g_\lambda} = [\mathbf{0.35}, \mathbf{0.35}, \mathbf{0.30}]$ # the weight of three metrics (SU, correlation_relevancy, correlation_redundancy]
3. Calculate Lambda using Eq.2 and $\boldsymbol{g_\lambda}$
4. **Repeat**
5. $\mathbf{f_i}$ ← feature i arrives at time i
6. $\boldsymbol{S_1}$=Calculate SU metric for feature $\mathbf{f_i}$ and class label Y using Eq.7
7. $\boldsymbol{S_2}$=Calculate correlation coefficient metric for feature $\mathbf{f_i}$ and class label Y using Eq.9
8. **if** S is empty, **then**
9. **if** $\max(\boldsymbol{S_1}, \boldsymbol{S_1}) > \mathbf{0.5}, \boldsymbol{then}\ \boldsymbol{S} = \boldsymbol{S} \cup \boldsymbol{f_i}$
10. **else** continue
11. **else**
12. $\boldsymbol{S_3}$=Calculate the average correlation coefficient metric for feature $\mathbf{f_i}$ and features in S using Eq.7
13. S3= -S3
14. Compute Sugeno-value using Eq.3 and Lambda
15. **if** Sugeno-value $> \mathbf{0.6}, \boldsymbol{then}\ \boldsymbol{S} = \boldsymbol{S} \cup \boldsymbol{f_i}$
16. **else** go to time i+1
17. **until** there is no new feature
18. **return** S

In this section, the description of algorithm 1 is presented. line1 refers to defining an empty vector (S) to store the final feature subset.

In combining values by Sugeno fuzzy integral method, we can assign different weights to each criterion that is equal to g_λ for each measure. Thus, in line 2, g_λ is determined for decision metrics. In our proposed algorithm, there are two types of feature evaluation metrics. The first type is relevancy measures. Therefore we have chosen two criteria (SU and Pearson correlation coefficient) for calculating the relevancy between each attribute and class label. Another criterion (Pearson correlation coefficient) is considered for calculating redundancy between the selected attribute and the new attribute as the second type of decision metric. We have assigned 0.35 to the relevancy criteria and 0.30 to the redundancy criterion as their weights in the aggregation process. Methods for evaluating features and their weights are experimentally adjusted. The number of relevancy criteria is higher because our goal is to give more importance to the relevance of the features.

In the third line of the algorithm, we calculate the lambda value based on Eq. 2, which is used throughout the algorithm to

calculate the Sugeno fuzzy integral value for each feature. Now we enter the feature selection phase, doing lines 4 to 18 of the algorithm iteratively for the number of input features. At any given time, a feature enters the decision-making space and needs to be decided at the moment whether accept or reject that feature. As a new attribute is entered, two relevancy values are calculated by Eq. 7 and Eq. 8. In the beginning, while no feature has not been selected, the concept of redundancy does not make sense. Therefore, we use the maximum value of two dependency criteria to select the first feature. In this case, if this maximum value is more significant than 0.5, the feature is selected. Otherwise, the algorithm waits for the nesting feature to enter. This process is performed in lines 8 to 10 of the algorithm.

After selecting the first feature, the algorithm enters the second phase. In this phase, a redundancy criterion is also calculated when a new feature arrives, in addition to relevancy metrics. For this matter, the correlation coefficient of the newly arrived features is computed against selected features. The average value is considered as our redundancy metric. We have three different values to decide whether a feature is selected or withdrawn. We chose the Sugeno fuzzy integral operator to aggregate these values using Eq. 3. To do this, we need the lambda that is computed before and $F(X)$. Considering that we have used three criteria, X=$\{S_1, S_2, S_3\}$ where S_1 to S_3 are our feature evaluation criteria obtained value. Thus, $F(S_1)$ to $F(S_3)$ are values that are computed for each feature using these metrics.

When the aggregation process is done, the acceptance or ignorance of the feature is conducted by the obtained value in this procedure. If this value is more significant than 0.6, then the feature is selected. Otherwise, the feature is ignored, and this process is repeated until no arriving feature is available. Two threshold values in our two phases are adjusted experimentally.

V. EXPERIMENTAL STUDIES

The performance of the proposed method is evaluated based on the comparison with five online streaming feature selection methods. These methods are OSFSMI [1], OFS-A3M [31], OFS-Density [24], SAOLA [19], and OFSS-FI [20] that use rank aggregation procedures.

A. Datasets

Six real-world datasets are considered to assess the performance of the proposed method and competitive methods. The description of these datasets is presented in Table I.

TABLE I. DESCRIPTION OF DATASETS

Dataset	Instances	Features	Classes	Type
Alon	62	2001	2	Microarray Data
ColonTumor	62	2000	2	Microarray Data
Prostate-GE	102	5967	2	Biological Data
SRBCT	82	2309	4	Microarray Data
Subramanian	50	10101	2	Image Data
West	49	7129	2	Microarray Data

B. Performance evaluation criteria

Accuracy and F1-score [32] measures are considered for measuring the classification performance of the proposed method and comparing methods. These metrics are defined based on the following concepts:

1) False-positive (FP): Percentage of negative examples incorrectly classified as positive.

2) True positive (TP): Percentage of positive examples correctly classified.

3) False-negative (FN): Percentage of positive examples incorrectly classified as negative.

4) True negative (TN): Percentage of negative examples correctly classified as so.

The accuracy and F1-score measures are defined as follows:

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{9}$$

$$F1 - score = \frac{\text{TP}}{\text{TP}+\frac{1}{2}(\text{FP+FN})} \tag{10}$$

C. Results

The recommended parameter values for OSFS methods are considered based on the corresponded paper in all competitive methods. The evaluation of classification accuracy for all comparing methods is performed based on the K-nearest neighbors (K-NN) and decision tree classifiers. The number of neighbors in the following classifier is set to 3 in all experiments.

We have set 60% of the instances in each dataset as training data randomly and the remaining 40% as the test data. 30 separate runs obtain the reported results of each method on every dataset. The presented values are the average value of the 30 runs.

Table II to Table V present the obtained classification results according to two classifiers and two classification metrics. In each table, the last row shows the average value for each algorithm on all datasets. The best performance between all algorithms in each dataset is bolded.

Also, in Table VI the average number of selected features by each measure is reported. The average run-time of all algorithms is also presented in Table VII.

TABLE II. RESULTS BASED ON KNN CLASSIFIER (ACCURACY)

dataset	OSFS MI	OFS-A3M	OFS-Density	SAOLA	OFSS-FI	OSFSSU
Alon	0.5958	0.6854	0.7188	0.6063	0.5563	**0.7646**
Colon	0.7438	0.7229	0.7563	0.7521	**0.7792**	0.7729
Prostate-GE	0.8300	0.7475	0.8988	0.8638	0.8575	**0.9263**
SRBCT	0.7136	0.8091	0.8318	0.7606	0.9015	**0.9273**
Subramanian	0.6150	0.6300	0.6175	0.5900	0.6525	**0.7050**
West	0.4842	0.6158	0.7395	0.6079	0.5500	**0.7579**
Mean	0.6637	0.7018	0.7604	0.6968	0.7162	**0.8090**

TABLE III. RESULTS BASED ON DECISION TREE CLASSIFIER (ACCURACY)

dataset	OSFS MI	OFS-A3M	OFS-Density	SAOLA	OFSS-FI	OSFSSU
Alon	0.5917	0.6542	0.7104	0.6313	0.6021	**0.7583**
Colon	0.7083	0.6750	0.7375	0.7167	0.7104	**0.7521**
Prostate-GE	0.8075	0.7788	**0.8650**	0.8350	0.8350	0.8175
SRBCT	0.7606	0.8182	0.8000	0.7621	0.7924	**0.8349**
Subramanian	0.5700	0.6075	0.5925	0.5900	0.6375	**0.6425**
West	0.5342	0.6184	0.6158	0.6684	0.6053	**0.7053**
Mean	0.6621	0.6920	0.7202	0.7006	0.6971	**0.7518**

TABLE IV. RESULTS BASED ON KNN CLASSIFIER (F-SCORE)

dataset	OSFS MI	OFS-A3M	OFS-Density	SAOLA	OFSS-FI	OSFSSU
Alon	0.3775	0.5056	0.5510	0.4160	0.3102	**0.6526**
Colon	0.7972	0.7776	0.8102	0.8016	**0.8316**	0.8226
Prostate-GE	0.8327	0.7370	0.8935	0.8617	0.8477	**0.9250**
SRBCT	0.6898	0.7974	0.8250	0.7535	0.9009	**0.9256**
Subramanian	0.7123	0.7152	0.7169	0.7047	0.7487	**0.7870**
West	0.5466	0.6759	0.7322	0.6417	0.5947	**0.7589**
Mean	0.6593	0.7014	0.7548	0.6965	0.7057	**0.8120**

TABLE V. RESULTS BASED ON DECISION TREE CLASSIFIER (F-SCORE)

dataset	OSFS MI	OFS-A3M	OFS-Density	SAOLA	OFSS-FI	OSFSSU
Alon	0.4635	0.5235	0.5736	0.5062	0.4338	**0.6321**
Colon	0.7789	0.7384	0.7978	0.7755	0.7591	**0.8154**
Prostate-GE	0.7988	0.7554	**0.8617**	0.8287	0.8250	0.8086
SRBCT	0.7799	0.8039	0.8104	0.7588	0.7980	**0.8394**
Subramanian	0.6601	0.6883	0.6858	0.6771	0.7141	**0.7345**
West	0.5301	0.6195	0.6227	0.6952	0.6107	**0.6877**
Mean	0.6685	0.6882	0.7253	0.7069	0.6901	**0.7530**

TABLE VI. THE AVERAGE NUMBER OF SELECTED FEATURES

dataset	OSFSMI	OFS-A3M	OFS-Density	SAOLA	OFSS-FI	OSFSSU
Alon	5	27.25	5.3	8.1	14.2	5.7
Colon	9.4	22.25	4.95	6.85	1921	6.75
Prostate-GE	9.25	38.9	6.55	7.55	186.95	24.95
SRBCT	12.45	11.2	4.5	5.7	2273.1	13.8
Subramanian	11.15	15.3	4.35	8.6	2436.3	9.15
West	12.6	24.4	3.85	8.95	1208.2	5.6
Mean	9.98	23.22	4.92	7.63	1339.9	10.99

TABLE VII. THE AVERAGE RUN-TIME OF OSFSSU AGAINST THE OSFS

dataset	OSFS MI	OFS-A3M	OFS-Density	SAOLA	OFSS-FI	OSFSSU
Alon	620.45	0.71	0.29	0.52	232.29	0.31
Colon	8.71	0.64	0.26	0.04	11.66	0.27
Prostate-GE	20.17	4.38	1.58	0.10	2.51	0.84
SRBCT	12.38	1.82	0.46	0.15	17.39	0.34
Subramanian	62.39	3.92	1.08	0.42	105.83	1.35
West	34.55	2.20	0.70	0.16	21.07	0.96
Mean	126.44	2.28	0.73	0.23	65.12	0.68

D. Discussion

OSFSSU method is a new algorithm for online streaming feature selection that uses the Sugeno fuzzy integral operator to combine multiple feature evaluation techniques. In this method, we have used a combination of redundancy-based and relevancy-based metrics to consider both the diversity and convergence in the proposed algorithm. This process causes the selection of less redundant and most relevant features subset. In the online streaming feature selection methods that have been proposed so far, the approach is based on the fact that new features are measured using a criterion. The feature is selected if the obtained value is above a pre-defined threshold. Otherwise, it is removed. In this paper, we present a new method that the measure for deciding on the acceptance of features, unlike current methods, is based on several criteria.

Since the variety of different methods can lead to a better result than one criterion. For this purpose, we have employed Sugeno fuzzy integral to combining the decision criteria. Sugeno fuzzy integral is used for this task because it can aggregate several values with a different effect on final results.

Table II to Table V shows that the proposed method is superior to competitive methods according to two classification measures and two different classifiers. Table VI shows the average number of selected features for each algorithm. We can see that the OSFSSU algorithm selects a small set of features, and at the same time, it has the best performance in the classification process. Also, the proposed method is performed quickly, and in just one case, the run-time is more than one second.

VI. CONCLUSION

This article proposes an online streaming feature selection method called the OSFSSU using Sugeno fuzzy integral. This paper calculated the value of three feature evaluation metrics for each new arriving feature. These values are delivered to the Sugeno fuzzy integral operator to perform the ensemble process. The results of different datasets indicate the efficiency and optimality of the proposed method since the convergence and diversity of the algorithm are balanced by the variety of feature selection methods. Also, due to the simple calculation in the process of the algorithm, the method performs in a short run-time, and the number of selected features is small. We intend to use ensemble techniques in other feature selection applications as our feature works and combine them with various optimization tasks in machine learning. Also, we want to improve our work in online streaming feature selection tasks and propose more powerful methods.

ACKNOWLEDGMENT

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

[1] M. Rahmaninia and P. Moradi, "OSFSMI: Online stream feature selection method based on mutual information," *Appl. Soft Comput.*, vol. 68, pp. 733–746, Jul. 2018, doi: 10.1016/j.asoc.2017.08.034.

[2] X. Hu, P. Zhou, P. Li, J. Wang, and X. Wu, "A survey on online feature selection with streaming features," *Front. Comput. Sci.*, vol. 12, no. 3, pp. 479–493, Jun. 2018, doi: 10.1007/s11704-016-5489-3.

[3] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-pour, "Ensemble of feature selection algorithms: a multi-criteria decision-making approach," *Int. J. Mach. Learn. Cybern.*, pp. 1–21, Jun. 2021, doi: 10.1007/s13042-021-01347-z.

[4] A. Hashemi and M. B. Dowlatshahi, "MLCR: A Fast Multi-label Feature Selection Method Based on K-means and L2-norm," in *2020 25th International Computer Conference, Computer Society of Iran (CSICC)*, Jan. 2020, pp. 1–7, doi:

10.1109/CSICC49403.2020.9050104.

[5] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-pour, "Gravitational Search Algorithm," in *Handbook of AI-based Metaheuristics*, 2021, p. 32.

[6] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-pour, "MFS-MCDM: Multi-label feature selection using multi-criteria decision making," *Knowledge-Based Syst.*, p. 106365, Aug. 2020, doi: 10.1016/j.knosys.2020.106365.

[7] S. R. Tiwari and K. K. Rana, "Feature Selection in Big Data: Trends and Challenges," 2021, pp. 83–98.

[8] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-Pour, "A bipartite matching-based feature selection for multi-label learning," *Int. J. Mach. Learn. Cybern.*, vol. 12, no. 2, pp. 459–475, Feb. 2021, doi: 10.1007/s13042-020-01180-w.

[9] A. Hashemi, M. Bagher Dowlatshahi, and H. Nezamabadi-pour, "An efficient Pareto-based feature selection algorithm for multi-label classification," *Inf. Sci. (Ny).*, vol. 581, pp. 428–447, Dec. 2021, doi: 10.1016/j.ins.2021.09.052.

[10] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-pour, "MGFS: A multi-label graph-based feature selection algorithm via PageRank centrality," *Expert Syst. Appl.*, vol. 142, p. 113024, Mar. 2020, doi: 10.1016/j.eswa.2019.113024.

[11] M. B. Dowlatshahi, V. Derhami, and H. Nezamabadi-pour, "Fuzzy particle swarm optimization with nearest-better neighborhood for multimodal optimization," *Iran. J. Fuzzy Syst.*, 2019, doi: 10.22111/ijfs.2019.5054.

[12] P. Dhal and C. Azad, "A comprehensive survey on feature selection in the various fields of machine learning," *Appl. Intell.*, Jul. 2021, doi: 10.1007/s10489-021-02550-9.

[13] A. Hashemi, M. Bagher Dowlatshahi, and H. Nezamabadi-pour, "VMFS: A VIKOR-based multi-target feature selection," *Expert Syst. Appl.*, p. 115224, May 2021, doi: 10.1016/j.eswa.2021.115224.

[14] H. Bayati, M. B. Dowlatshahi, and M. Paniri, "MLPSO: A Filter Multi-label Feature Selection Based on Particle Swarm Optimization," in *2020 25th International Computer Conference, Computer Society of Iran (CSICC)*, Jan. 2020, pp. 1–6, doi: 10.1109/CSICC49403.2020.9050087.

[15] Jialei Wang, Peilin Zhao, S. C. H. Hoi, and Rong Jin, "Online Feature Selection and Its Applications," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 3, pp. 698–710, Mar. 2014, doi: 10.1109/TKDE.2013.32.

[16] F. BenSaid and A. M. Alimi, "Online feature selection system for big data classification based on multi-objective automated negotiation," *Pattern Recognit.*, vol. 110, p. 107629, Feb. 2021, doi: 10.1016/j.patcog.2020.107629.

[17] M. Paniri, M. B. Dowlatshahi, and H. Nezamabadi-pour, "Ant-TD: Ant colony optimization plus temporal difference reinforcement learning for multi-label feature selection," *Swarm Evol. Comput.*, vol. 64, p. 100892, Jul. 2021, doi: 10.1016/j.swevo.2021.100892.

[18] P. Zhou, P. Li, S. Zhao, and Y. Zhang, "Online early terminated streaming feature selection based on Rough Set theory," *Appl. Soft Comput.*, vol. 113, p. 107993, Dec. 2021, doi: 10.1016/j.asoc.2021.107993.

[19] K. Yu, X. Wu, W. Ding, and J. Pei, "Towards Scalable and Accurate Online Feature Selection for Big Data," in *2014 IEEE International Conference on Data Mining*, Dec. 2014, pp. 660–669, doi: 10.1109/ICDM.2014.63.

[20] P. Zhou, P. Li, S. Zhao, and X. Wu, "Feature Interaction for Streaming Feature Selection," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 32, no. 10, pp. 4691–4702, Oct. 2021, doi: 10.1109/TNNLS.2020.3025922.

[21] J. Zhou, D. P. Foster, R. A. Stine, and L. H. Ungar, "Streamwise feature selection," *J. Mach. Learn. Res.*, vol. 3, no. 2, pp. 1532–4435, 2006.

[22] P. Zhou, X. Hu, P. Li, and X. Wu, "Online feature selection for high-dimensional class-imbalanced data," *Knowledge-Based Syst.*, vol. 136, pp. 187–199, Nov. 2017, doi: 10.1016/j.knosys.2017.09.006.

[23] P. Zhou, X. Hu, P. Li, and X. Wu, "Online streaming feature selection using adapted Neighborhood Rough Set," *Inf. Sci. (Ny).*, vol. 481, pp. 258–279, May 2019, doi: 10.1016/j.ins.2018.12.074.

[24] P. Zhou, X. Hu, P. Li, and X. Wu, "OFS-Density: A novel online streaming feature selection method," *Pattern Recognit.*, vol. 86, pp. 48–61, Feb. 2019, doi: 10.1016/j.patcog.2018.08.009.

[25] G. Beliakov and D. Divakov, "On representation of fuzzy measures for learning Choquet and Sugeno integrals," *Knowledge-Based Syst.*, vol. 189, p. 105134, Feb. 2020, doi: 10.1016/j.knosys.2019.105134.

[26] M. Ayub, "Choquet and Sugeno Integrals," 2009.

[27] J.-L. Marichal, "On Sugeno integral as an aggregation function," *Fuzzy Sets Syst.*, vol. 114, no. 3, pp. 347–365, Sep. 2000, doi: 10.1016/S0165-0114(98)00116-X.

[28] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-Pour, "Minimum redundancy maximum relevance ensemble feature selection: A bi-objective Pareto-based approach," *J. Soft Comput. Inf. Technol.*, 2021.

[29] S. Kashef and H. Nezamabadi-pour, "A label-specific multi-label feature selection algorithm based on the Pareto dominance concept," *Pattern Recognit.*, vol. 88, pp. 654–667, 2019, doi: 10.1016/j.patcog.2018.12.020.

[30] M. C. ABOUNAIMA, F. Z. EL MAZOURI, L. LAMRINI, N. NFISSI, N. EL MAKHFI, and M. OUZARF, "The Pearson Correlation Coefficient Applied to Compare Multi-Criteria Methods: Case the Ranking Problematic," in *2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, Apr. 2020, pp. 1–6, doi: 10.1109/IRASET48871.2020.9092242.

[31] P. Drotár, M. Gazda, and L. Vokorokos, "Ensemble feature selection using election methods and ranker clustering," *Inf. Sci. (Ny).*, vol. 480, pp. 365–380, 2019, doi: 10.1016/j.ins.2018.12.033.

[32] A. Hashemi, M. Bagher Dowlatshahi, and H. Nezamabadi-pour, "A pareto-based ensemble of feature selection algorithms," *Expert Syst. Appl.*, vol. 180, p. 115130, Oct. 2021, doi: 10.1016/j.eswa.2021.115130.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Power Allocation in CRNs based on QoS and QoE

Marjaneh Moghbeli Damaneh[1], Omid Abedi[2], Mahdieh Ghazvini[3], Mostafa Ghazizadeh Ahsaee[4]

[1]*Department of Computer Engineering, Shahid Bahonar University of Kerman, Kerman, Iran,* marjmoghbeli@gmail.com
[2]*Department of Computer Engineering, Shahid Bahonar University of Kerman, Kerman, Iran,* oabedi@uk.ac.ir
[3]*Department of Computer Engineering, Shahid Bahonar University of Kerman, Kerman, Iran,* mghazvini@uk.ac.ir
[4]*Department of Computer Engineering, Shahid Bahonar University of Kerman, Kerman, Iran,* mghazizadeh@uk.ac.ir

***Abstract*— In Cognitive Radio Networks, secondary users have access to the channel Opportunistically. So, it must be ensured that the severe interference, that damages the main network, does not occur. Using the resource allocation methods provides the best access to the spectrum for the secondary users, increases their performance, and primary users are protected from collisions. The Underlay method is one of the methods of the Cognitive Radio Network. In this method the secondary and the primary users simultaneously present in the channel. The power allocation in an underlay Cognitive Radio Network, may ensure that the secondary users do not interfere with the primary users. Using optimization techniques is a way to improve the performance of these networks. The NSGA-III algorithm, a multi-Objective optimization method, is efficient and offers acceptable results. This paper presents a power allocation method in the Underlay Cognitive Radio Networks based on the improvement of QoS and QoE, using the NSGA-III algorithm.**

***Index Terms* —Cognitive Radio Networks, Underlay, Power Allocation, QoS, QoE, NSGA-III.**

I. INTRODUCTION

Cognitive radio networks (CRN) have introduced as a promising solution to the problem of spectrum shortages and inefficiency of wireless networks [1-2]. This technology makes optimal use of available spectral resources and reduces interference with other users. Primary users (PU) are the main users of the spectrum and the owner of the spectrum who can use the spectrum whenever they want, and secondary users (SU) are users who are let to use the spectrum due to network conditions and the absence of primary users. In CRNs, the spectrum can be shared in three scenarios: interweave, overlay, and underlay [3]. In the Interweave scenario, the spectrum can only be operated by SU if it is not used by PU. As a result, spectrum measurement is essential in this mode, and incorrect spectrum measurement can reduce network performance [4-5]. In an overlay scenario, there is a collaboration between the main network and the secondary network, the use of spectrum by SUs is also done in small intervals, where PUs are not present. A SU adjusts its transmission characteristics to avoid interference [6-7]. In the underlay scenario, SUs can access the spectrum simultaneously with the PUs, provided there is no interference with the PUs. This mode can increase spectrum Efficiency if proper spectrum management and resource allocation techniques are used.

Resource allocation (RA) is a crucial function in wireless communications. Resources include time slots, frequency bands, transmission power, TXRX antennas, and so on. RA plans are appropriate for better utilization of available resources. In CRN, the SUs achieve the spectrum opportunistically, so no interference must be guaranteed. The use of RA algorithms to supply the SUs with the best spectral gain is such that it enhances their efficiency while protecting the system from collisions. RA improves Quality of Service (QoS) and Quality of Experience (QoE). QoS include latency, vibration, bandwidth, throughput, etc, that describe network conditions. QoE is a general acceptance of services by the user such as Mean Opinion Score (MOS), PSNR, etc. QoE is also relevant to user settings as well as some immeasurable and subjective parameters such as users' capabilities, expectations, and feelings [8-9]. Providing the same QoS level creates other satisfaction levels for different users and varies over time [10]. In all the mentioned articles, other methods have been used to achieve the goals, other ways in allocating power.

Genetic algorithm (GA) is a population-based processes for optimization and search problems that uses biological techniques such as inheritance and mutation. The engine of the genetic algorithm creates an initial population of formulas. Each person is evaluated against a set of data, and after each generation, the most appropriate ones remain, and the rest are discarded. Genetic algorithms tend to create precise formulas

after many generations. In 1995, the Non-dominated Sorting Genetic Algorithm (NSGA) was introduced to solve multi-Objective optimization (MO) problems. It has also been shown [13] that among the various MO genetic algorithms, NSGA offers much better results. Owing to the high sensibility of the efficiency and modality of the NSGA algorithm answers to match parameters, the NSGA-II algorithm was proposed by. In addition to all the performance of the NSGA-II meta-heuristic algorithm, it can be considered as a model for the formation of many MO algorithms [14]. This algorithm and its unique approach to MO problems have been used repeatedly by different people to create newer MO algorithms. In [15], the NSGA-III algorithm is presented for problems with a large number of objectives (three to 15 objectives) and is better than other algorithms and provides satisfactory results for all issues considered. The NSGA-III algorithm uses the NSGA-II algorithm, which is used to achieve many goals.

In this paper, a power allocation method for underlay cognitive radio networks based on MOS, fairness, and energy consumption goals using the NSGA-III algorithm is presented. In Section 2, we describe the related works. Section 3 discusses the proposed method. In Section 4, the simulation results are reviewed, and finally, in Section 5, conclusions are made.

II. RELATED WORKS

In CRNs, maximizing one criterion usually minimizes other network metrics, so researchers have recently achieved balanced performance in the network by considering MOP problems. Related works can be categorized from two perspectives of power allocation based on QoS and QoE. For example, in [16], the problem of user grouping and power allocation are provided that different groups are separated by the OMA method, and users in a group have NOMA access. The key idea of NOMA is to search the power domain for multiple accesses so that users in the same group can access the spectrum at the same time. The purpose is to maximize throughput according to power constraints. To ensure fairness, the certain weight is set for each group, which is multiplied by the rate. Weight is determined by channel conditions so that users with poor channel conditions can be less stressed and can transmit with a higher power. The grouping problem is solved with the Distributed Sequential Coalition Formation (DSCF) method and the power allocation problem is solved with the Swarm Intelligence algorithm (SI). In [17], the power allocation scheme for underlay CR users is introduced. Collision control and collision power limitation play an important role in CRNs, the convex optimization theory is applied to the power allocation for the maximum transmission rate (channel capacity) under the three constraints of the collision, total power, and power in time-varying channels. Also, in [18], the allocation of resources for the underlay spectrum in CRNs has been investigated. There is a limit to the collision with PUs and a QoS limit for SUs, and the collision of SUs and PUs must be below the threshold. An acceptance control algorithm for rate and power control with max-min fairness and proportional fairness is considered. The goal is maximizing rate for fairness and all restrictions. The stepwise maximum interference removal algorithm (SMIRA) algorithm has been investigated to remove the minimum number of secondary links due to QoS constraints and PU limitations. This algorithm is combined with interference constraints (I-SMIRA). The new algorithm has a lower probability of failure and a higher number of accepted links. In [19] a robust energy efficiency control algorithm for CRNs is introduced. The power has a threshold limit, and SINR has different limits over time. The purpose of is to maximize energy efficiency. To protect the PUs from a collision when the power does not exceed the threshold, a protection factor has been introduced, which can be considered as a variable in time due to the limitations of the SINR variable over time. There is also another factor for SINR limitation that varies over time (protection limit for SINR users). Based on these two constraints, the Robust Power Algorithm (REEA) is introduced, which allocates less power and SINR to SU users and thus improves energy efficiency. In [20], The purpose of this paper is to allocate fairly resources by reducing channel collision (CCI). The function of utility is the probability of correct acceptance (CRP), meaning that the signal sent in the channel is transmitted and received correctly. Power allocation for a channel with maximum power limit is performed with Perron Frobenius nonlinear theory. In [21], The performance of genetic algorithms in power allocation for subscript CRNs is investigated. Power allocation plays an important role in performance improvement of CRNs, since it has to not only limit the interference with PUs, but also meet the minimum QoS requirements of SUs with the confined resources. The goal is the maximum probability of simultaneous transfer, the maximum performance (capacity) of the SUs, and the maximum energy efficiency (in contrast to the energy consumption) under the conditions of meeting the minimum QoS for the SU and PU. To increase the performance or rate, the transmission power must be increased, and conversely, as the transmission power decreases, the energy efficiency improved. All goals depend on the transmission power and distance between PU and SU users. In [22], GA-based optimal power allocation for underlay CRNs has been performed. Two Utility functions are considered: 1) Maximizing system capacity by considering power limit, SINR, and collision power, 2) minimize power consumption. In [23], the selection of optimal flexural power has been performed in opportunistic underlay CRNs. Improving the operational capacity and QoS and overcome collisions. The goal is to reduce the possibility of misdiagnosis and false alarms. In some studies, power allocation is based on the QoE, for example, in [24], the goal is maximizing MOS in the combined services for video streaming and file download programs with a limited bit error rate,

Table I. SUMMARY OF RELATED WORKS

Reference	Objective	Method used	Advantages	Disadvantages
[16]	Maximum throughput	DSCF Algorithm for Spectrum Allocation and Power Allocation with SI Algorithm.	Ensure dealing with PU and fairness, get maximum throughput.	Ignoring delay, and energy consumption.
[17]	Maximum transmission rate	Power allocation using convex optimization theory	Guaranteed SU collision with PU and maximum transmission rate.	Ignoring latency, fairness, and energy consumption.
[18]	Maximum transfer rate and fairness.	I-SMIRA acceptance control algorithm for power allocation.	Guarantee of non-collision, Guarantee of QoS in SU and consideration of fairness and maximum transfer rate.	Ignoring latency, fairness and energy consumption.
[19]	Maximum energy efficiency.	REEA Algorithm for power control.	Guarantee of non-collision and improving energy efficiency.	Ignoring transfer rates, fairness, and other parameters.
[20]	Maximum Accurate Acceptance Probability	Using Perron Frobenius nonlinear theory for power allocation.	Reduce channel collisions and consider fairness and maximum CRP.	Ignoring transmission rates, energy consumption and delays.
[21]	Maximum probability of simultaneous transmission, capacity of SUs and energy efficiency.	Using genetic algorithms for power allocation.	Collision power limit, QoS guarantee of SU and PU, maximum simultaneous transmission probability, SU capacity and energy efficiency.	Ignoring priority, and fairness between SUs, and delays.
[22]	Maximum system capacity, minimum power consumption.	Optimal power allocation with genetic algorithm.	Collision limitation and protection of PUs, maximum system capacity with minimum power consumption.	Ignorance of QoS, latency, and fairness.
[23]	Improved throughput and QoS, reduced detection errors.	Use of a flexible power transmission method, power control circuit design.	Limit dealing with PU, improve throughput and QoS, reduce detection errors.	Ignoring fairness, energy consumption, and delay.
[24]	The Maximum MOS and transfer rate.	Use geometric programming and convex optimization for power allocation.	PU collision limit, MOS maximum and transmission rate.	Ignoring energy consumption, and delays.
[25]	The Maximum MOS about fairness.	Use heuristic algorithm for channel allocation and Fast Barrier method for power allocation.	Limit dealing with PU, maximum MOS and fairness and increase the overall rate of SUs.	Ignoring energy consumption and delay.
[26]	The Maximum MOS.	Use DQN and learn transitions to allocate power.	PU interference limit, MOS maximum, convergence time reduction.	Ignoring energy consumption, and delay.
[27]	Reduce convergence time.	Use collaborative learning to allocate power.	Limit interference with PU, reduce convergence time.	Ignoring latency, fairness, and energy consumption.

maximum power, and collision limit. In [25], Orthogonal Frequency Division Multiplexing (OFDM) in used in the CR transmission power and collision limit, with optimal channel allocation and power. In [26], The goal is to reach the best balance between minimizing the detrimental effects of SUs and maximizing QoE using Deep Q-Learning Network (DQN) and transfer learning.

Similar work has also been done to reduce convergence time by learning to cooperate with power allocation [27]. A summary of the methods presented in Table 1.

III. THE PROPOSED METHOD

In the underlay CRNs, the absence of collisions between PUs and SUs is the main condition of communication. SINR indicates the SINR; the higher the SINR, the better the audio and video quality. Here it is assumed that the PU transmits at a constant power. SINR for PU at the base station (PBS), $SINR^{(P)}$ and for the SU_i at its corresponding SBS, $SINR^{(S)}$, may be represented as follows: [27]

$$SINR^{(P)} = \frac{G_0^P P_0}{\sigma^2 + \sum_{j=1}^{N} G_j^{(P)} P_j}, \tag{1}$$

$$SINR_i^{(s)} = \frac{G_i^S P_i}{\sigma^2 + G_0^S P_0 + \sum_{j \neq i} G_j^{(s)} P_j}, \tag{2}$$

P_0 is the transfer function of PU, and P_i is the transfer function of SU_j. G_0^P and $G_j^{(P)}$ is the PU to PBS channel gains and SU_j's channel gains to the PBS. QoE imposes limitations on the minimum allowable SINRs of the network; the SINR threshold is applied by introducing β_0 and β_i:

$$\begin{cases} SINR^{(P)} \geq \beta_0 \\ SINR_i^{(s)} \geq \beta_i \ , \qquad i = 1, \dots, N. \end{cases} \tag{3}$$

By introducing SINR limitations for primary and secondary networks, the power control solution for each SU is as follows: [28]

$$\begin{cases} P_i = \frac{\Psi_i(\sigma^2 + G_0^S P_0)}{G_i^{(s)}(1 - \sum_{j=1}^{N} \Psi_j)}, \qquad i = 1, \dots, N, \\ \Psi_i = (1 + \frac{1}{\beta_i})^{-1} \end{cases} \tag{4}$$

The condition $1-\sum_{j=1}^{N}\Psi_j>0$ requires a valid power allocation. Substituting the SU powers obtained from (4) into (1), (3), the following relations can be rewritten:

$$\sum_{j=1}^{N}\alpha_j\Psi_j\leq 1, \tag{5}$$

Where

$$\alpha_j=\frac{G_j^{(p)}(\sigma^2+G_0^{(s)}P_0)}{G_j^{(s)}(G_0^{(p)}P_0/\beta_0-\sigma^2)}+1. \tag{6}$$

from (4) and (5), it is obvious that $\beta_i$must be tuned in each SU to not only to meet the needs of dynamic spectrum access interference but also to increase the overall QoE of the network. The goal is to acquire, the best balance between minimizing the detrimental effects of SUs and maximizing QoE. To set the intended system, the relationship between the transmission bit rate and the SINR of the SU threshold is as follows [29]:

$$r_i^{(s)}=Wlog_2(1+k\beta_i) \tag{7}$$

k= $\frac{1.5}{-\ln(5BER)}$ is a constant that depends on the value of the maximum transmission bit error rate (BER). The task of optimization is to maximize the performance of the network. QoE attracts a lot of attention as a user representative. MOS has been selected as one of the most widely used network performance metrics. For data traffic, the MOS is estimated according to the transmission rate $r_i^{(s)}$, experienced by the end-user, as follows [30]:

$$Q_D=alog_{10}(br_i^{(s)}(1-P_{e2e})), \tag{8}$$

Q_D and P_{e2e} are data MOS and packet loss factors. while parameters a, and b are computed by the maximum and minimum sufficient data realized by end-users. If the transmission rate of a user is R and the efficient receiving rate is R, then the probability of packet destruction is zero; The perceived quality rate of the end-user based on MOS might be maximum, i.e., 5. In addition, the value of MOS, 1 is assigned to the minimum transfer rate. In this work $a=1.3619$ and $b=0.6780$. For video traffic, objective quality is often calculated based on the peak signal-to-noise ratio (PSNR), because in video traffic, as opposed to data traffic, the target is video quality. It does not matter if we have a small number of lost packets. However, PSNR does not exactly understand the human mental perception of video quality. In [31], a logic function is used to determine the relationship between MOS and PSNR as follows:

$$Q_v=\frac{c}{1+\exp(d(PSNR-h))}, \tag{9}$$

Where $Q_v$represents the MOS for the video, and a, b and c are logical function parameters. In the function $PSNR=kr_i^{(s)}+p$, $r_i^{(s)}$is transmission bit rate and k and p are constants. In this work, k = 10.4 and p = 28.7221, c = 6.6431, d = 0.1344 and h = 30.4264.

All users in cognitive radio networks should have a fair opportunity to access resources, and resources should be shared fairly between users. Also, energy consumption should be fair, and user's QoS and QoE needs should be relatively proportionate [32]. Lack of fairness leads to a lack of QoS in the network. fairness is defined according to the Jain Index as follows [33]:

$$Fairness(S)=\frac{(\sum_{i=1}^{N}QOE_i(S))^2}{N\sum_{i=1}^{N}(QOE_i(S))^2} \tag{10}$$

N is the number of secondary users in the cognitive radio network. Maintaining fairness among secondary users is another goal of this article. The goal is maximizing MOS of all users if it increases the fairness. The combination of MOS and fairness parameters is done as follows:

$$MOS(S)=\sum_{i=1}^{N}MOS_i(S). \tag{11}$$

$$\mathrm{M(S)}=Fairness(S)\times\frac{MOS(S)}{N} \tag{12}$$

MOS (S) is the maximum MOS for all users, and Fairness (S) is fairness for all secondary users. By combining these two parameters, a new parameter M (S) is obtained, which aims to maximize this parameter [10]. Another goal of this article is to get the maximum transfer rate.

$$Max\sum_{k=1}^{K}R_K \tag{13}$$

One approach to increasing network capacity or rates is increasing the transmission capacity of SUs, which leads to power consumption and energy consumption. As a result, minimizing this energy consumption, is an issue that needs to be addressed, and minimizing it is one of this article's goals.

$$\mathrm{Min}\ \frac{\sum_{i=1}^{i}P_i}{\sum_{i=1}^{i}R_i} \tag{14}$$

The NSGA-III algorithm quantifies a random population of size N to find the optimal solutions. It calculates the objective functions and then sorts the population by rank and dispersion distance. it uses genetic operators (selection, crossover, and mutation) to obtain new children. The population combines parents and children to form the new population. On the selected answers to change, a point is randomly chosen on the chromosome string and then moved to the left or right areas of that point in the chromosome string (intersection). Newborns enter the new generation with the best of the previous generation, then the best is selected in each generation, and this cycle is repeated, and if the value of the function does not change, the best chromosome in the last generation is presented as the answer to the problem (mutation). These steps are performed repeatedly to achieve optimal goals. The set of solutions here are the same capabilities that can be assigned to SUs. The pseudo-code of the proposed method is shown below.

Algorithm NSGA-III for Power Allocation

```
for time it <MaxIt do
  for all SU_i,i=1,...,N do
     Random generation of initial population
     Calculate objective functions. (MOS, Fairness, Power Consumption)
     Non-dominated Sorting Based on rank and crowding distance
     Selection, Applying crossover and mutation for Generation
     Population composition of parents and children
  end for
end for
```

Table II. SIMULATION PARAMETERS

Parameter	Value
Size of Population	80
Generations Number	50
Probability of Crossover	0.5

Optimization algorithms must be repeated regularly to achieve optimal results. the number of iterations that an optimization algorithm needs to achieve optimal results is essential. Getting the best products at the cost of high convergence repetitions is not worth it. In this paper, we will have a comparison of the number of convergence iterations in the NSGA-III algorithm and other algorithms.

I. SIMULATION RESULTS

An initial network contains an initial link to access a single channel. The SINR objective for PU is adjusted to 1dB. Gaussian noise power and PU transmission power are 1 and 10 watts. All SUs and PUs are distributed randomly around their corresponding BSs in a circle with a radius of 300 m. For this simulation we consider the number of SUs {4,8,12,16,20}. Simulation parameters are given in Table II. We present the results of the NSGA-III algorithm in three different modes: 1. When the target is the MOS criterion. 2. When the target is the MOS and Fairness criteria. 3. When the target is the MOS criteria, Fairness, and energy consumption, for different number of SUs, we review and compare with the simple Q-learning algorithm and the improved Q-learning in [27]. Deciding on these algorithms and obtaining optimization results is based on these three goals. Fig. 1 compares MOS in simple Q-Learning, improved Q-Learning, and NSGA-III in three different models: (a) considering MOS, (b) considering MOS and fairness, (c)) Considering MOS, Fairness, and Power. As can be seen, the NSGA-III algorithm has a higher MOS value in all cases than the other two algorithms. In Fig. 2 compares bit rate of the three algorithms in three different modes. As can be seen, the NSGA-III algorithm has a higher bit rate value in all cases than the other two algorithms. Fig. 3 also compares the energy consumption of the three algorithms in three different modes. As shown in the figure, The NSGA-III algorithm consumes almost less power than the other two algorithms. Fig. 4 also compares number of convergence iterations of the three algorithms in three different modes, As shown, NSGA-III algorithm has a relatively good convergence. The NSGA-III algorithm has achieved better results than the simple Q-Learning algorithm and the improved Q-Learning algorithm in all three different modes due to selecting the best solutions.

I. CONCLUSION

In this paper, we examine the allocation of power in an underlay cognitive radio network considering maximum MOS and fairness considering energy consumption. To have a better system, we must balance these parameters by evaluating the limitation of interference with PUs and ensuring a minimum QoS in SUs in power allocation. The NSGA-III algorithm is beneficial for MO problems and provides acceptable results. The simulation results illustrate that the proposed method gives much better results in allocating MO power than the Q-Learning algorithm and increases the QoS and the QoE.

REFERENCES

[1] M. Amjad, M. H. Rehmani, and S. Mao, "Wireless multimedia cognitive radio networks: A comprehensive survey," IEEE Communications Surveys & Tutorials, vol. 20, no. 2, pp. 1056-1103, 2018.

[2] M. A. Shah, S. Zhang, and C. Maple, "Cognitive radio networks for Internet of Things: Applications, challenges and future," in 2013 19th International Conference on Automation and Computing, 2013, pp. 1-6: IEEE.

[3] M. El Tanab and W. Hamouda, "Resource allocation for underlay cognitive radio networks: A survey," IEEE Communications Surveys & Tutorials, vol. 19, no. 2, pp. 1249-1276, 2016.

[4] O. Altrad, S. Muhaidat, A. Al-Dweik, A. Shami, and P. D. Yoo, "Opportunistic spectrum access in cognitive radio networks under imperfect spectrum sensing," IEEE Transactions on Vehicular Technology, vol. 63, no. 2, pp. 920-925, 2013.

[5] A. Ali and W. Hamouda, "Low power wideband sensing for one-bit quantized cognitive radio systems," IEEE Wireless Communications Letters, vol. 5, no. 1, pp. 16-19, 2015.

[6] J. Hua and N. Jiang, "QoS performance analysis for the second user in the overlay cognitive radio networks," in 2015 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery, 2015, pp. 513-516: IEEE.

[7] W. Yifei, T. Yinglei, W. Li, S. Mei, and W. Xiaojun, "QoS Provisioning energy saving dynamic access policy for overlay cognitive radio networks with hidden Markov channels," China Communications, vol. 10, no. 12, pp. 92-101, 2013.

[8] J. Zhang and N. Ansari, "On assuring end-to-end QoE in next generation networks: challenges and a possible solution," IEEE Communications Magazine, vol. 49, no. 7, pp. 185-191, 2011.

[9] R. Stankiewicz and A. Jajszczyk, "A survey of QoE assurance in converged networks," Computer Networks, vol. 55, no. 7, pp. 1459-1473, 2011.

[10] B. S. Ghahfarokhi, "Distributed QoE-aware channel assignment algorithms for IEEE 802.11 WLANs," Wireless Networks, vol. 21, no. 1, pp. 21-34, 2015.

[11] N. Srinivas and K. Deb, "Muiltiobjective optimization using nondominated sorting in genetic algorithms," Evolutionary computation, vol. 2, no. 3, pp. 221-248, 1994.

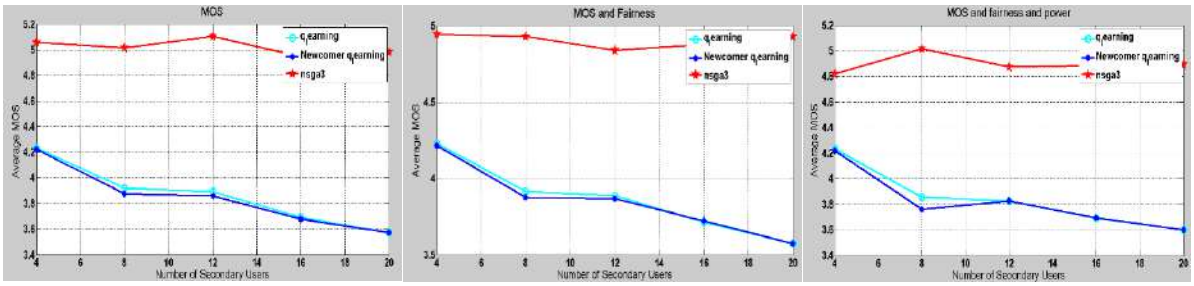
[12] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," IEEE transactions on evolutionary computation, vol. 6, no. 2, pp. 182-197, 2002.

[13] A. H. Dias and J. A. De Vasconcelos, "Multiobjective genetic algorithms applied to solve optimization problems," IEEE Transactions on Magnetics, vol. 38, no. 2, pp. 1133-1136, 2002.

[14] B. Gadhvi, V. Savsani, and V. Patel, "Multi-objective optimization of vehicle passive suspension system using NSGA-II, SPEA2 and PESA-II," Procedia Technology, vol. 23, pp. 361-368, 2016.

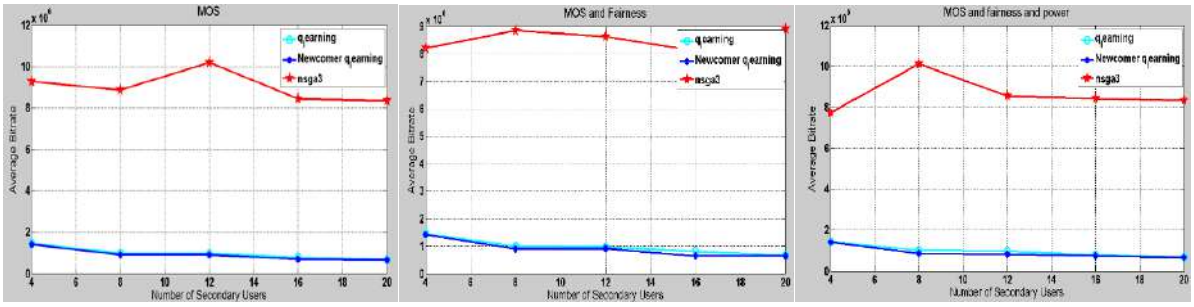
[15] K. Deb and H. Jain, "An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: solving problems with box constraints," IEEE Transactions on Evolutionary Computation, vol. 18, no. 4, pp. 577-601, 2013.

[16] W. Liang, K. Di Wang, J. Shi, L. Li, and G. K. Karagiannidis, "Distributed Sequential Coalition Formation Algorithm for Spectrum Allocation in Underlay Cognitive Radio Networks," IEEE Access, vol. 7, pp. 56803-56816, 2019.



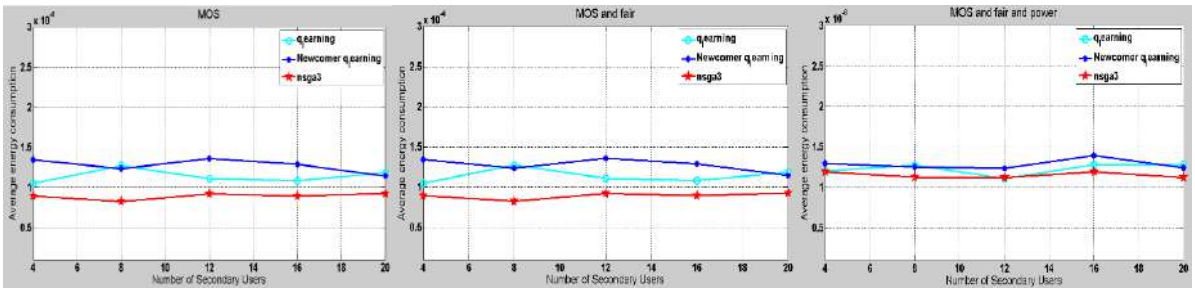
(a) (b) (c)

Figure 1. Comparison of MOS in Simple Q-Learning, Improved Q-Learning and NSGA-III in three different models: (a) Considering MOS, (b) Considering MOS and fairness, (c) Considering MOS, Fairness, and Power.



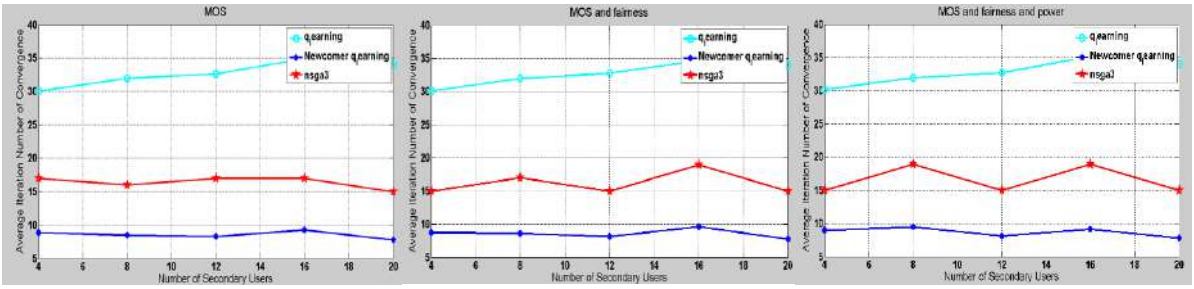
(a) (b) (c)

Figure 2. Comparison of bit rate in simple Q-Learning, improved Q-Learning and NSGA-III in three different models: (a) considering MOS, (b) considering MOS and fairness, (c)) Considering MOS, Fairness, and Power.



(a) (b) (c)

Figure 3. Comparison of energy consumption in simple Q-Learning, improved Q-Learning and NSGA-III in three different models: (a) considering MOS, (b) considering MOS and fairness, (c)) Considering MOS, Fairness, and Power.



(a) (b) (c)

Figure 4. Comparison of the number of convergence iterations in simple Q-Learning, improved Q-Learning and NSGA-III in three different models: (a) considering MOS, (b) considering MOS and fairness, (c)) Considering MOS, Fairness, and Power.

[17] Y. Li, G. Zhu, S. Lin, K. Guan, L. Liu, and Y. Li, "Power allocation scheme for the underlay cognitive radio user with the imperfect channel state information," 2013.

[18] L. B. Le and E. Hossain, "Resource allocation for spectrum underlay in cognitive radio networks," IEEE Transactions on Wireless communications, vol. 7, no. 12, pp. 5306-5315, 2008.

[19] M. Zhou, X. Zhao, and H. Yin, "A robust energy-efficient power control algorithm for cognitive radio networks," Wireless Networks, pp. 1-10, 2017.

[20] W. Xu, R. Qiu, and J. Cheng, "Fair Optimal Resource Allocation in Cognitive Radio Networks With Co-Channel Interference Mitigation," IEEE Access, vol. 6, pp. 37418-37429, 2018.

[21] D. Bepari and D. Mitra, "Performance of GA in Power Allocation for Underlay Cognitive Radio Systems," Journal of Communications Technology and Electronics, vol. 63, no. 10, pp. 1250-1258, 2018.

[22] D. Bepari and D. Mitra, "GA based optimal power allocation for underlay cognitive radio networks," in 2014 International Conference on Electronics and Communication Systems (ICECS), 2014, pp. 1-6: IEEE.

[23] S. Chatterjee, P. Banerjee, and M. Nasipuri, "Optimized Flexible Power Selection for Opportunistic Underlay Cognitive Radio Networks," Wireless Personal Communications, vol. 96, no. 1, pp. 1193-1213, 2017.

[24] B. Liu and L. He, "QoE-based resource allocation for mixed services over cognitive radio networks," in 2016 International Conference on Computing, Networking and Communications (ICNC), 2016, pp. 1-5: IEEE.

[25] J. Dai and S. Wang, "QoE-driven resource allocation method for cognitive radio networks," in 2016 IEEE International Conference on Communications (ICC), 2016, pp. 1-6: IEEE.

[26] F. Shah-Mohammadi and A. Kwasinski, "Deep Reinforcement Learning Approach to QoE-Driven Resource Allocation for Spectrum Underlay in Cognitive Radio Networks," in 2018 IEEE International Conference on Communications Workshops (ICC Workshops), 2018, pp. 1-6: IEEE.

[27] F. S. Mohammadi and A. Kwasinski, "QoE-Driven Integrated Heterogeneous Traffic Resource Allocation Based on Cooperative Learning for 5G Cognitive Radio Networks," in 2018 IEEE 5G World Forum (5GWF), 2018, pp. 244-249: IEEE.

[28] S. Pietrzyk and G. J. Janssen, "Radio resource allocation for cellular networks based on OFDMA with QoS guarantees," in IEEE Global Telecommunications Conference, 2004. GLOBECOM'04., 2004, vol. 4, pp. 2694-2699: IEEE.

[29] X. Qiu and K. Chawla, "On the performance of adaptive modulation in cellular systems," IEEE transactions on Communications, vol. 47, no. 6, pp. 884-895, 1999.

[30] O. Dobrijevic, A. J. Kassler, L. Skorin-Kapov, and M. Matijasevic, "Q-point: QoE-driven path optimization model for multimedia services," in International Conference on Wired/Wireless Internet Communications, 2014, pp. 134-147: Springer.

[31] P. Hanhart and T. Ebrahimi, "Calculation of average coding efficiency based on subjective quality scores," Journal of Visual Communication and Image Representation, vol. 25, no. 3, pp. 555-564, 2014.

[32] S. Huaizhou, R. V. Prasad, E. Onur, and I. Niemegeers, "Fairness in wireless networks: Issues, measures and challenges," IEEE Communications Surveys & Tutorials, vol. 16, no. 1, pp. 5-24, 2013.

[33] R. Jain, The art of computer systems performance analysis: techniques for experimental design, measurement, simulation, and modeling. John Wiley & Sons, 1990.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Reinforcement Learning Reward Function for Test Case Prioritization in Continuous Integration

Hajar Mirzaei[1], Mohammad Reza Keyvanpour[2]

[1]*Department of Computer Engineering, Faculty of Engineering, Alzahra University, Tehran, Iran*, hajarmirzaei248@gmail.com
[2] *Department of Computer Engineering, Faculty of Engineering, Alzahra University, Tehran, Iran*, keyvanpour@alzahra.ac.ir

***Abstract*— Given that software systems are constantly changing at a fast rate during the software development process, Continuous integration testing that a cost-effective software development practice is characterized by constantly evolving test cases, rapid return, and limited performance time. Regression test is performed after any change in the software, so we must seek to optimize the regression test methods. One of the methods that have attracted a lot of attention today is the prioritization of the test using reinforcement learning. The aim of this study is to review and compare reinforcement learning reward functions, which are used to test case prioritization. We divided these reward functions into two categories: reward functions that depend on current information and reward functions that depend on historical information. The reward functions based on current information rely only on the information of a previous implementation and the historical information-based reward functions rely on the whole information of the past. Each of these functions has the strengths and weaknesses that we have discussed in this study. Among the reward functions of these two categories, the HFCW reward function has better performance both in terms of fault detection and usage time.**

***Index Terms* —Continuous integration, Test Case Prioritization, Reinforcement learning, Machine Learning, Reward function**

I. INTRODUCTION

Modern software systems are constantly changing at fast rate during the software development process. Changes are made to the software to improve performance, add new features and fix identified software bugs [1]. These changes to the software may alter the performance of the software or may result in the software to be working bizarre or may affects various other parts of the software, Therefore, in order to prevent this, regression testing have been performed [2]. Regression testing (RS) is an important activity to guarantee software quality. Regression testing ensures that new changes in the software will not affect the performance of the current software. Software researchers have paid special attention to Continuous Integration (CI) in recent years. CI is a cost-effective software development practice commonly where developers often combine their work [3]. It has a number of tasks, containing software configuration management, version control, auto-build, and RT of candidates for the new software version [4]. Due to the complication of modern software systems, it is very important to maintain reliability and quality, both economically and economically, particularly in large, rapidly evolving enterprises [5]. Many industries are using CI strategy for this reason. CI strategy is a beloved software development technique in which developers often merge the latest corrections to their code, via a commit, into the main code database, allowing them to cost-effectively and easily verify which their code can successfully pass tests in different system surroundings [4].

The most important techniques for the optimization of regression testing are Test Case Selection (TCS), Test Suite Minimization (TSM), and Test Case Prioritization (TCP). The test case selection only performs test cases which traverse through the corrected portions of the system-under-test. Although the test case selection temporarily follows the changes introduced, test suite minimization permanently removes test cases. TCS, meticulous selection of regression test cases based on coverage is required. Therefore, suitability coupled with crucial test-redundancy jointly plays a major role in securing a TSM [4]. The purpose of TCP techniques is to reorganize test cases to maximize early failure detection. This approach has the advantage over TCS and TSM in that it does not rule out any test cases from the implementation. TCP methods provide a way to perform test cases with higher fault detection capability earlier to provide early feedback to developers. TCP also makes it possible to continue the test up to the time or budget limit, by executing the test suite in the order specified by prioritization. The TCP, which is the subject of this study, has been extensively

studied and numerous approaches have been suggested for the TCP problem.

Machine learning (ML) has already been used in TCP in 2006, but over the past few years, it has become more popular in TCP fields and researchers have been relying more and more on ML techniques to conduct TCP. These studies showed promising results for new techniques, reinforcing ML's view as an effective alternative to traditional TCP approaches. Reinforcement learning (RL) is a subset of ML that is good position to develop an adaptive approach that can learn from its experience in the performance environment. Adaptive means that our method can gradually increase its effectiveness based on observations of the effects of its actions. In this paper, we review the performance of RL in the TCP process.

The document's structure is as follows: Sect. 2 sets out the background and related work. It also in Sect. 2 referred to main and fundamental concepts such as test case prioritization, machine learning, reinforcement learning for test case prioritization. Sect. 3 presents a reward for TCP based on RL. Sect. 4 we evaluate and analyze the rewards discussed in this article. Lastly, Sect. 5 concludes the article.

II. BACKGROUND AND RELATED WORK

A. Test Case Prioritization

The purpose of prioritizing test cases is to detect a regular sequence of test cases in order to maximize efficiency measures. In many literary works, TCP algorithms goal to uncover failures as soon as possible [11]. In more formal terms, Rothermel et al. [5] describe the TCP problem in the following manner.

TCP given a test suite T and the set of its all possible permutations PT, and a function from PT to real numbers $f : PT \rightarrow \mathbb{R}$, find a proper subset $\acute{T}$ so that

$$[f(T') \geq f(T'')], \quad \forall T'' \in PT \qquad (1)$$

Within this definition, PT is the set of all potential prioritization of T and f is any function that determines the efficiency of a given prioritization.

During the prioritization of tests, testers can reach various aims by maximizing performance metrics or certain criteria. The definition of performance can be different because testers will have various aims at various times. Testers may first want to achieve maximal code coverage; they may later want to find as many faults as possible. According to these scenarios, defining the TCP problem is the identical, however, the performance function changes under optimization. Therefore, there are multiple prioritization criteria, such as code coverage, maximizing the probability of exposing faults early, program changes, the probability of disclosing errors associated with particular code changes, and test case diversity [11] [12].

B. Machine Learning

Test approaches based on ML are usually based on source code information. Catal [18] studied these fault-prediction methods and showed that ML techniques can improve the likelihood of fault detection in comparison to classic software reviews utilizing specific code criteria[19]. In recent years, researchers have been utilized machine learning-based techniques for conducting test case prioritization. By using learning algorithms, the system can utilize the learning from prior tests and adapt to evolution applications. Test cases define the priority with continuous actions and machine learning techniques offer easy adaptation to changes [20].

The main concept of using these techniques implies the creation of an educational dataset based on different characteristics of the test cases (Test cases are usually prioritized based on various factors or artifacts, such as cost-awareness, coverage information, requirements, and specification, etc.). Afterward, the test cases are labeled with several priority classes based on specific criteria. A machine learning algorithm is then trained on the training dataset. After training, it can classify new test cases according to their priority classes. Finally, the test cases predicted with the highest priority are executed earlier, revealing the most important faults early in the testing phase.

As shown in Figure 1, based on the applied ML techniques, the studies are classified by the machine learning technique into four groups: Natural Language Processing-based (NLP-based) Models, Reinforcement Learning (RL), Ranking Models (RM) and Clustering (CL) [4]. In the following, we examine the reinforcement learning that is the main subject of this paper.

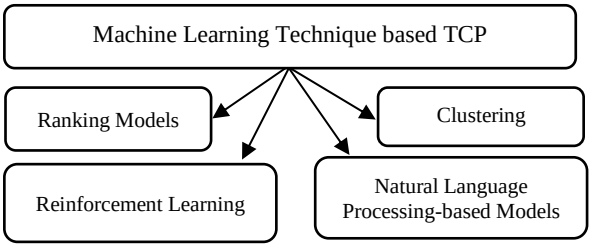


Figure 1. Machine Learning Technique based TCP

C. Reinforcement Learning for TCP

Reinforcement learning, as one of the important sub-branches of machine learning, is a cyclical process that predicts an artificial factor that learns from the environment by observing its situation and selecting an appropriate action, given a set of possible actions and states [21]. As a result of the acts, the agent receives feedback in the form of a reward, and every action executed has a specific reward. The goal of a factor (determiner), is to interface with the surroundings and learn in a 'trial and error' manner, and choose actions that maximize the total accumulation of the reward signal [5,6]. The value of a negative or positive reward is calculated based on to the feedback of history behavior by a reward function, and utilized for the next action that the agent decides. Determining the suitable reward value may have a considerable impact on the result of reinforcement learning. If the reward function is not properly crafted, it is hard to learn the intended target for the agent [10].

RL approach can be divided into two classes: model-free and model-based. Model-free, the approach does not have a basic concept of the dynamics of the environment and the way its

actions effect it [4]. Indeed, in the model-free approach, algorithms rather than models are mostly utilized to optimise for performance. On the other hand, the model-based approach has a problem domain model which is constructed and utilized in the process of selecting actions [13].

In the TCP framework, each test case is exclusively prioritized and then test program is created, performed, and eventually evaluated.

Figure 2 shows the linkages between TCP and RL. Each status represents the metadata of a single test case, comprising information on the duration of the test case, the time of the most recent performed, and the prior performance results. As an action, the test case's priority by the RL agent assigned for the current CI cycle is returned. After all test cases are prioritized, the prioritized test suite is planned, inclusive of a selection of the more major test cases, and presented for implementation and finally, their respective rewards are ascribed according to the situation of the test case for feedback and performance assessment. Based on this reward, the agent may match their strategy for future actions [4][5].

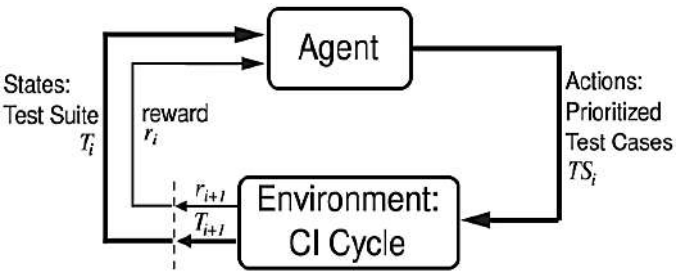


Figure 2. Interaction of Agent and Environment [5]

III. Reward Function for RL in CI Testing

Reward functions are first studied to measure the failure detection capacity of a test case in continuous integration tests. These reward functions are divided into two groups: the group that uses historical information and the group that use current information. In the group that uses historical information, using all historical information about test cases. But, for the group using the current information, only the performance result of the latest generation of test cases is utilized to calculate the reward value of the reward function [14]. However, it is difficult to specify whether a test case should be rewarded only according to the outcome of current implementation result because the historical results are more valuable to assess the failure detection capacity for a test case [14]. Figure 3 shows the RL reward function classification.

A. Failure Count Reward Function (FC)

In the FC, every test cases, receive the count of unsuccessful test cases from the test sequence in every cycle as a reward. The FC is an intuitive reward function that rewards the RL agent directly in order to maximize the count of unsuccessful test cases [11]. The definition is as follows:

$$FC_i(t) = \left|T_i^{fail}\right| \qquad (\forall\, t \in T_i) \qquad (2)$$

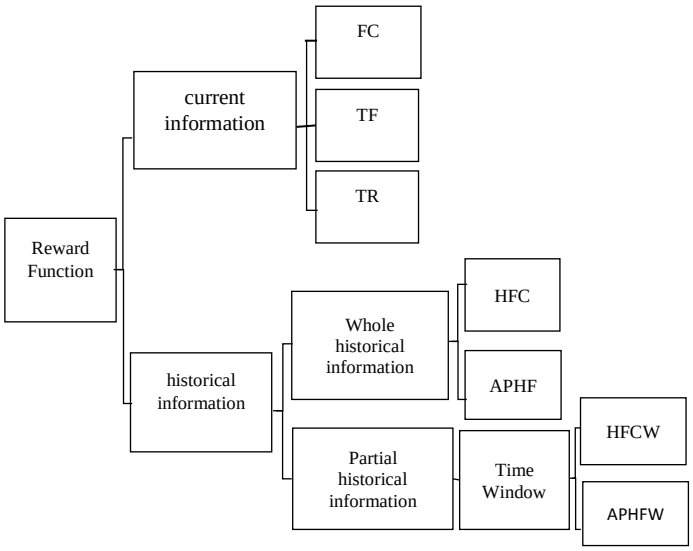


Figure 3. The frame of reward Function.

B. Test Case Failure Reward Function (TF)

TF rewards only the test case that failed in the current cycle and brings back the verdict of the test case as a single reward for any test case. Planning for test failures is intended and consequently strengthened. If a test case passes, no particular reward is given as inclusive, nor reduces or improves the quality of the schedule based on the information available. However, the sequence of test cases is not specifically included within the reward. This is implicit in encouraging the agent to focus on prioritizing test cases and failing them above.

$$TF_i(t) = \begin{cases} 1 - t \cdot verdict_i & if\ t \in T_i \\ 0 & otherwise \end{cases} \qquad (3)$$

C. Time-ranked Reward Function(TR)

TR comprises the sequence of test cases and rewards each test case for ranking on the test program and if it is unsuccessful. Because an awesome calendar performs failed test cases early, each passed test case decreases the quality of the calendar if it precedes a failed test case. Every test case is rewarded with a total number of test cases unsuccessful for test cases unsuccessful TR is identical to FC. For passing test cases, the TR is further reduced by the count of test cases unsuccessful classified after the successful test case, which penalizes early test case planning [11].

$$TR_i(t) = \left|T_i^{fail}\right| - t * s_i \times \sum 1, \qquad (4)$$
$$t_k \in T_i^{fail} \wedge rank(t) < rank(t_k)$$

In Equations (2), (3) and (4), T is a set of test cases$\{t_1, t_2, \ldots, t_n\}$. A few of these test cases are chosen and requested for implementation called T_i ($T_i \subset T$) . When resources or time constraints do not exist, T_i may be defined to hold the orderly sequence of all test cases, T_i^{tot}. When such constraints apply, a sub-set of the T test pool is chosen for T_i. Notice that T is an unordered set, while both T_i and T_i^{tot} are

ordered sequences. Thus, should be defined a ranking function that acts over all test cases: rank: $T_i \rightarrow N$, where $rank(t)$ is the position of t within T_i. In T_i, every test case t has a duration $t.d_i$, that is known before performance, and a status, $t.s_i$, which is known after performance, that is equal to 0 if the test has failed, or 1 if it has passed. In T_i, the sub-set of all failed test cases is noted $T_i^{fail} = \{t \in T_i \, s.t. \, t.s_i = 0\}$, and $\left|T_i^{fail}\right|$ is the count of test cases failures.

D. Historical Failure Count Reward Function (HFC)

In the CI test process, the historical implementation information for a test case is the visualization of the failure detection capability. The more errors a test case is detected, the greater the likelihood it is that the next integration will detect the errors. Therefore, the count of failures of a test case can be utilized to measure the failure detection capability of the test case, which can be utilized as a reward [14].

In this document, rwd_j is used to report the result of executing the test case in its j_{th} implementation. rwd_j values are 0 or 1, where 0 indicates pass and 1 indicates failure. The definition of the HFC reward function is as follows:

$$HFC_i(t) = \sum_{j=1}^{n} rwd_j \tag{5}$$

Where i means the i_{th} integration, n means the count of historical implementations of the test case t. It is important to note that a test case is not necessarily performed in each integration, therefore the i_{th}execution does not show the j_{th}CI cycle.

E. Average Percentage of Historical Failure Rreward Function (APHF)

The APHF is defined to measure the distribution of historical faults in a test case, whereby the APHF is defined as follows:

$$APHF_i(t) = 1 - \frac{\sum_{i=1}^{m} R_j}{n \times m} + \frac{1}{2n} \tag{6}$$

Where i means the i_{th} integration, n means the count of historical implementations of the test case t, and m means the failure number of t in n implementations. R_j means the countd0wn order of the last j_{th} failure of t.

The APHF range is from 0 to 1. The higher the APHF value, the higher probability of failure of test case. Therefore, the test cases with a higher APHF value are more probably to discover an error in the next implementation [14].

F. Reward Function Based on Time-Window

Feedback time should be as brief as conceivable because time is important in CI testing. In comparison with the reward function based on current information, it takes longer in computing the reward function based on historical information [14]. Thus, it is unnecessary to consider all historical information because some of them, which has not been used for a long time, may become obsolete. On the other hand, the recent history is of greater importance than early testing for test case sorting techniques. As a result, a time window-based method is used in the reward function to extract more valuable historical information from recent test cases in the earlier cases, indeed the reward function uses a time window to filter recent historical information to calculate reward value [10]. According to the definition of HFC and APHF, the reward function based on time-window, called HFCW and APHFW, and is defined in Equations (7) and (8), respectively.

- *Historical Failure Count with Time Window (HFCW).* HFCW computes the count of historical failures for test case t in prior w implementation, in which w is the time window size and it is defined as:

$$HFCW_i(t) = \sum_{j=1}^{w} rwd_j \tag{7}$$

- *Average Percentage of Historical Failure with time Window (APHFW).* APHFW calculates the average percentage of historical failure for test case t in the prior w implementation, and it is defined as:

$$APHFW_i(t) = 1 - \frac{\sum_{i=1}^{m} R_j}{win(n,w) \times m} + \frac{1}{2 \times win(n,w)} \tag{8}$$

Where $win(n,w)$ denotes the count of historical implementations of test case t in the time window w, and m is the count of failed implementations of $win(n,w)$ [14].

$$win(n,w) = \begin{cases} n & n < w \\ m & n \geq w \end{cases} \tag{9}$$

IV. EVALUATION

Numerous researches have been conducted in Reinforcement Learning based Test Case Prioritization. In this article, in order to evaluate the reward functions, previous works have been used. In this section, a comparison of reward functions is presented. During the first part, an overview of the evaluation metric used in the experiment was given and then the results were analyzed and discussed.

A. Evaluation Metrics

To compare the performance of various approaches, evaluation criteria are needed as a common performance index. Normalized Average Percentage of Faults Detected (NAPFD) [16] is a common utilized to assess prioritization test cases. For the calculation of NAPFD, the ratio of discovered and detectable faults in T is taken into account. This is enough to prioritization of test cases when not all of them are performed, and faults may not be discovered [17]. Equation 10 describes how to calculate NAPFD.

$$NAPFD(T_i) = P - \frac{\sum_{j \in T_i^{fail}} rank(j)}{\left|T_i^{fail}\right| \times |T_i|} + \frac{P}{2 \times |T_i|} \tag{10}$$

With $P = \frac{\left|T_i^{fail}\right|}{\left|T_i^{tot,fail}\right|}$, $rank(j)$ indicates the position of the j_{th} failed test case in the test sequence T_i, $\left|T_i^{fail}\right|$represents the total

number of test cases failed in T_i, $|T_i|$ represents the total number of test cases in T_i, and $\left|T_i^{tot,fail}\right|$ represents the total number of test cases failed in the T.

Time Consumption (TC) is the average time cost of reinforced test case selection. This covers time spent calculating reward function, reinforcement learning, and sorting test cases [14].

B. Analyzed And Discussed

The TCP attempts to set the test case which can discover faults as much as conceivable in the front. The reward function in reinforcement learning assesses the test case's fault detection capability with the reward value. In this section, we discuss the efficiency of various reward functions using the NAPFD and CT assessment metrics, and the results of this analysis are presented in Table I.

TABLE I. THE REWARD FUNCTION ANALYSIS USING EVALUATION METRICS

	Current Information			**Historical Information**			
				Whole Historical		**Partial Historical Time-Window**	
	FC	**TF**	**TR**	**HFC**	**APHF**	**HFCW**	**APHFW**
NAPFD	Med	Low	Low	High	High	High	High
TC	Low	Low	Low	High	High	Med	Low

Spieker et al. [5] presented and tested FC, TR, TF reward functions. Spieker's experimental findings indicated that for test cases of failed, the TR reward function is the same as the FC reward function. For test cases passed, the TR reward function is further reduced by the count of test cases of failed after the test case passed in order to penalize the scheduling of early passing test cases. In addition, conclude that TF's test sequence performed better than TR and FC.

According to the experiments of Yang et al. [14] for historical information-based reward functions, HFC that rewards a test case based on its number of unsuccessful historical implementations, to some extent it may well reflect the defect detection capability of a test case. Nonetheless, for test cases with an identical number of faults, HFC rewards have the identical ability to detect faults, whereas test cases with newer failures are most likely to discover failures. In this manner, by increasing CI cycles, the failure data of a test case tends to have less impact on the assessment of fault discovery capacity.

For instance, test cases $t_1$and t_2 are both performed n times, and the results of their historical implementation are shown as $[1, 0, \ldots, 0]$ and $[0, \ldots, 0, 1]$. One can see that the count of failures of t_1 and t_2 have the identical value of 1 in terms of HFC reward, but t_1 unsuccessful in the last implementation, and t_2 unsuccessful in the first implementation.

Utilizing the APHF, the reward values to t_1 and t_2 are $1 - \frac{1}{2n}$ and $\frac{1}{2n}$, respectively. While t_1 and t_2 have the same number of failures, t_1 detects a fault in the last implementation which has a greater APHF value, that is, a greater likelihood of identifying defects in the next merger.

A more detailed examination reveals that with the increase of n, the value of the APHF (t_2) approaches infinitely 0. Therefore, as CI cycles increase, the failure data of a test case tends to have less effect on the evaluation of identifying defect capacity. The APHF can represent this tendency correctly.

With respect to the reward function based on time window, we have come to the conclusion that the value of the HFCW is no greater than the value of the HFC, Duo to the decrease in the count of historic failures. Therefore, HFCW concentrates on the effect of the latest failures, test case information on fault detection.

The APHFW varies between 0 and 1. Since APHFW also focuses on the effect of the recent time window failures rather than the full historical implementations, the computing cost may be lower than that of APHF. Notice which for a test case that does not fail in the time window w performances, the reward is reduced to 0, which can affect the assessment of the test case [14].

It is important to note that, we must determine a suitable range of $win(n,W)$, several time window sizes may result in different results. It should be noted that when W is small, which containing TF ($W = 1$), the fault detection capacity is low at a little time cost. When W is greater, it means that it takes into account all performance history, and APHFW has better fault discovery performance at a higher time cost. In reality, APHF ($W = n$)and TF ($W = 1$) are the specific forms of APHFW (W from 1 to n). Generally, APHFWs that use recent historical data have better outcomes than TF and do not have lower performance than APHF. This means that APHFW believes that historical information is sufficient, but not superfluous, to establish a reasonable prioritization list, that can discover fault faster. In the time window-based reward functions, in addition to saving time, it also improves the NAPFD average [11].

Given the above, the reward function based on the historical information is more appropriate for low failure systems, which more accurately reflects the actual position of software development. The current information-based reward function is most appropriate when software development usually fails early. But the reward functions based on time window combines the salient feature of reward functions based on current information and historical information, and achieve the best performance in all data sets [14].

V. CONCLUSION

In this study, we are focused on reinforcement learning reward functions for test case prioritization in Continuous integration. We examined and compared reward functions in terms of information. And we came to the conclusion that in current information-based reward functions, the FC reward function performs better. In historical information-based reward functions and FC, the APHFW reward function performed better both in terms of time and error detection.

REFERENCES

[1] A. Kumar, "Development at the Speed and Scale of Google. QCon San Francisco," ed, 2010.

[2] A. Memon et al., "Taming Google-scale continuous testing," in 2017 IEEE/ACM 39th International Conference on Software Engineering: Software Engineering in Practice Track (ICSE-SEIP), 2017: IEEE, pp. 233-242.

[3] E. Albert, M. Gómez-Zamalloa, and G. Puebla, "PET: a partial evaluation-based test case generation tool for Java bytecode," in Proceedings of the 2010 ACM SIGPLAN workshop on Partial evaluation and program manipulation, 2010, pp. 25-28.

[4] R. Pan, M. Bagherzadeh, T. A. Ghaleb, and L. Briand, "Test case selection and prioritization using machine learning: a systematic literature review," Empirical Software Engineering, vol. 27, no. 2, pp. 1-43, 2022.

[5] H. Spieker, A. Gotlieb, D. Marijan, and M. Mossige, "Reinforcement learning for automatic test case prioritization and selection in continuous integration," in Proceedings of the 26th ACM SIGSOFT International Symposium on Software Testing and Analysis, 2017, pp. 12-22.

[6] M. Bagherzadeh, N. Kahani, and L. Briand, "Reinforcement learning for test case prioritization," IEEE Transactions on Software Engineering, 2021.

[7] T. Shi, L. Xiao, and K. Wu, "Reinforcement Learning Based Test Case Prioritization for Enhancing the Security of Software," in 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA), 2020: IEEE, pp. 663-672.

[8] S. Mondal and R. Nasre, "Hansie: Hybrid and consensus regression test prioritization," Journal of Systems and Software, vol. 172, p. 110850, 2021.

[9] G. Rothermel, R. H. Untch, C. Chu, and M. J. Harrold, "Prioritizing test cases for regression testing," IEEE Transactions on software engineering, vol. 27, no. 10, pp. 929-948, 2001.

[10] M. Mahdieh, S.-H. Mirian-Hosseinabadi, K. Etemadi, A. Nosrati, and S. Jalali, "Incorporating fault-proneness estimations into coverage-based test case prioritization methods," Information and Software Technology, vol. 121, p. 106269, 2020.

[11] Z. Wu, Y. Yang, Z. Li, and R. Zhao, "A time window based reinforcement learning reward for test case prioritization in continuous integration," in Proceedings of the 11th Asia-Pacific Symposium on Internetware, 2019, pp. 1-6.

[12] A. Nguyen, B. Le, and V. Nguyen, "Prioritizing automated user interface tests using reinforcement learning," in Proceedings of the Fifteenth International Conference on Predictive Models and Data Analytics in Software Engineering, 2019, pp. 56-65.

[13] V. Nguyen and B. Le, "RLTCP: A reinforcement learning approach to prioritizing automated user interface tests," Information and Software Technology, vol. 136, p. 106574, 2021.

[14] Y. Yang, Z. Li, L. He, and R. Zhao, "A systematic study of reward for reinforcement learning based continuous integration testing," Journal of Systems and Software, vol. 170, p. 110787, 2020.

[15] J. Campos, A. Arcuri, G. Fraser, and R. Abreu, "Continuous test generation: Enhancing continuous integration with automated test generation," in Proceedings of the 29th ACM/IEEE international conference on Automated software engineering, 2014, pp. 55-66.

[16] X. Qu, M. B. Cohen, and K. M. Woolf, "Combinatorial interaction regression testing: A study of test case generation and prioritization," in 2007 IEEE International Conference on Software Maintenance, 2007: IEEE, pp. 255-264.

[17] J. A. P. Lima and S. R. Vergilio, "Test Case Prioritization in Continuous Integration environments: A systematic mapping study," Information and Software Technology, vol. 121, p. 106268, 2020.

[18] C. Catal, "Software fault prediction: A literature review and current trends," Expert systems with applications, vol. 38, no. 4, pp. 4626-4636, 2011.

[19] R. Lachmann, S. Schulze, M. Nieke, C. Seidl, and I. Schaefer, "System-level test case prioritization using machine learning," in 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), 2016: IEEE, pp. 361-368.

[20] E. K. Meçe, H. Paci, and K. Binjaku, "The Application Of Machine Learning In Test Case Prioritization-A Review," European Journal of Electrical Engineering and Computer Science, vol. 4, no. 1, 2020.

[21] A. Bertolino, A. Guerriero, B. Miranda, R. Pietrantuono, and S. Russo, "Learning-to-rank vs ranking-to-learn: strategies for regression testing in continuous integration," in Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering, 2020, pp. 1-12.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Residuated lattice congruences via directed kernels

1st Farideh Frasad, 2nd Arsham Borumand Saeid

1st Department of Pure Mathematics, Faculty of Mathematics and Computer, Higher Education Complex of Bam, Bam, Iran,
faridehfarsad@yahoo.com
2nd Department of Pure Mathematics, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman, Kerman, Iran,
arsham@uk.ac.ir

***Abstract*— To study the quotient of algebras, like residuated lattices, whose algebraic structures is determined by a partial order, it is often more common to think about directed kernels of homomorphisms between such algebras. So, in this paper, we first introduce the concept of a pre-congruence on a residuated lattice and then characterize them as the directed kernels of the residuated lattice homomorphisms. Second, we characterize the residuated lattice congruences as the intersection of a pre-congruence and its inverse.**

***Index Terms*— Pre-congruence, Residuated lattice congruence, Residuated lattice, quotient.**

I. Introduction and Preliminaries

The interest in lattice-valued logic has been rapidly growing recently. Several algebras playing the role of structures of true values have been introduced and axiomatized [1–3]. The most general structure considered in this paper is that of a residuated lattice.

Residuated lattices were first introduced as a generalization of ideal lattices of rings in 1939 by Ward and Dilworth[5].

For residuated lattices, whose algebraic structure is determined by a partial order, it is more useful to consider the sub-kernels of residuated lattice homomorphisms to study the residuated lattice congruences. So, in this paper, we take a close look at in sub-kernels of residuated lattice homomorphisms and characterize them as pre-congruences on a residuated lattices. Moreover, we characterize residuated lattice congruences as the intersection of a sub-kernel relation and its inverse.

Let's now recall for example from [5] the notion of residuated lattices and residuated lattice homomorphisms.

Definition 1. *A residuated lattice is an algebra* $(P, \wedge, \vee, \odot, \rightarrow, 0, 1)$ *of type (2,2,2,2, 0,0) equipped with an order* $\leq$ *satisfying the following:*
(LR_1) $(P, \wedge, \vee, 0, 1)$ *is a bounded lattice,*
(LR_2) $(P, \odot, 1)$ *is a commutative ordered monoid,*
(LR_3) $\odot$ *and* $\rightarrow$ *form an adjoint pair, i.e.,* $c \leq a \rightarrow b$ *iff* $a \odot c \leq b$, *for all* $a, b, c \in P$.

The relations between the pair of operations $\odot$ and $\rightarrow$ expressed by LR_3, is a particular case of the law of residuation, or Galois correspondence (see [3]) and for every $x, y \in A$, $x \rightarrow y = \sup\{z \in A \mid x \odot z \leq y\}$.

Definition 2. *A function* $f : P \rightarrow Q$ *between two residuated lattices* P *and* Q *is called a residuated lattice homomorphism if it is a morphism of bounded lattices and for every* $x, y \in P$: $f(x \odot y) = f(x) \odot f(y)$ *and* $f(x \rightarrow y) = f(x) \rightarrow f(y)$.

Example 1. *If on* $I = [0, 1]$; *for* $x, y \in I$ *we define* $x \odot y = \min\{x, y\}$ *and* $x \rightarrow y = 1$ *if* $x \leq y$ *and* y *otherwise, then* $(I, \max, \min, \odot, \rightarrow 0, 1)$ *is a residuated. lattice.*

II. Pre-complete pre-ordered sets

First of all, we recall some concepts of pre-ordered sets which will be used in the sequel (see [3]).
Consider a set P with a reflexive and transitive relation σ. Such a relation will be called a *pre-order* and (P, σ) *a pre-ordered set*.

In the following, we give an example of a pre-ordered set which is not a poset.

Example 2. *Let* X *be a set with* $|X| \geq 2$, *and consider the total relation* $\leq = X \times X$ *on it. Then* $(X, \leq)$ *is a pre-order. In fact, for every* $A \subseteq X$, *each element of* X *is a pre-supremum and pre-infimum of it. But* $(X, \leq)$ *is not a poset, because the binary relation* $\leq$ *is not a partial order.*

Let (P, σ) be a pre-ordered set. Then the relation $\overline{\sigma}$ defined by $p \ \overline{\sigma} \ q$ if and only if $p \ \sigma \ q$ and $q \ \sigma \ p$, for each $p, q \in P$, is an equivalence relation on P. Moreover, it is the largest equivalence relation on P with the property that $\overline{\sigma} \subseteq \sigma$.

It is a well-known fact that $(P/\overline{\sigma}, \sigma/\overline{\sigma})$ is a poset. Moreover, $\overline{\sigma}$ is the smallest equivalence relation with this property. We also have that for every $X, Y \in P/\overline{\sigma}$, $X \ (\sigma/\overline{\sigma}) \ Y$ if and only if $x \ \sigma \ y$ for some $x \in X$ and $y \in Y$ if and only if $x \ \sigma \ y$ for each $x \in X$ and $y \in Y$.

Definition 3. *Let $(P, \leq)$ be a pre-ordered set and let $A \subseteq P$. An element $x \in P$ is a pre-upper bound of A if $a \leq x$, for all $a \in A$. A pre-lower bound is defined dually. In other words, an element $y \in P$ is a pre-lower bound of A if $y \leq a$, for all $a \in A$.*

Definition 4. *Let $(P, \leq)$ be a pre-ordered set and let $A \subseteq P$. An element $x \in P$ is called a pre-supremum of A if*

(i) *x is a pre-upper bound of A, and*
(ii) *$x \leq y$ for all pre-upper bounds y of A.*

A pre-infimum of A is also defined dually (see Example 2).

Remark 1. *Since $\leq$ is not necessarily anti-symmetric, a subset A of P may have more than one a pre-supremum and pre-infimum. But for every pair x, x' of pre-suprema (pre-infima), we have $x \leq x' \leq x$ or equivalently $x \overline{\leq} x'$. The set of all pre-suprema and the set of all pre-infima of A are denoted by Pre-sup(A) and Pre-inf(A), respectively.*

Lemma 1. *Let P be a non-empty pre-ordered set such that Pre-sup$(A) \neq \emptyset$, for all $A \subseteq P$. Then Pre-inf$(A) \neq \emptyset$, for all $A \subseteq P$.*

Proof: Let $A \subseteq P$ and let A' be the set of all pre-lower bounds of A. Hence, by assumption, Pre-sup$(A') \neq \emptyset$. We claim that Pre-sup(A') = Pre-inf(A). First notice that every element of A is a pre-upper bound of all pre-infima of A. Hence every element $b \in$ Pre-sup(A') is a pre-lower bound of A. Further, if c is any pre-lower bound of A then $c \in A'$ and so $c \leq b$ ($b \in$ Pre-sup(A')). This implies that Pre-sup$(A') \subseteq$ Pre-inf(A). Conversely, if z is a pre-infimum of A, then $z \in A'$ and also $t \leq z$ for all $z \in A'$. This gives that z is a pre-upper bound of A'. Now, if w is any pre-upper bound of A' then $z \leq w$ ($z \in A'$). This implies that $z \in$ Pre-sup(A') and so Pre-inf$(A) \subseteq$ Pre-sup(A'), as required ∎

Theorem 1. *Let P be a non-empty pre-ordered set. Then the following are equivalent:*

(i) *P is pre-complete.*
(ii) *Pre-sup$(A) \neq \emptyset$, for all $A \subseteq P$.*
(iii) *Pre-inf$(A) \neq \emptyset$, for all $A \subseteq P$.*

Proof: It is trivial that (i) implies (ii), and ,by Lemma 1, (ii) implies (iii). By dually of the proof of the above lemma (iii) implies that (i). ∎

III. Pre-congruence relations on Residuated lattices

In the following, we define a pre-congruence relation on a residuated lattice which has an important role in characterizing residuated lattice congruences.

Definition 5. *Let P and Q be two pre-complete pre-ordered sets and $f : P \to Q$ be a map.*

(i) *f preserves arbitrary pre-suprema if s is a pre-supremum of $A \subseteq P$ then $f(s)$ is a pre-supremum of $f(A)$ in Q, for all $A \subseteq P$.*
(ii) *f preserves arbitrary pre-infima if s is a pre-infimum of $A \subseteq P$ then $f(s)$ is a pre-infimum of $f(A)$ in Q, for all $A \subseteq P$.*

Definition 6. *Let $(P, \leq)$ be a residuated lattice, a pre-order σ on P is called pre-congruence if it satisfies the following conditions:*

(1) *$\leq \subseteq \sigma$.*
(2) *(P, σ) has binary pre-suprema and binary pre-infima and $(P, \sigma, \odot)$ is a pomonoid.*
(3) *$\odot$ and $\to$ form an adjoint pair, i.e., $c\ \sigma\ (a \to b)$ iff $(a \odot c)\ \sigma\ b$, for all $a, b, c \in A$, where $a \to b =$ pre-suprema$\{c \in P \mid (a \odot c)\ \sigma\ b\}$.*
(4) *The identity function $id_P : (P, \leq) \longrightarrow (P, \sigma)$ preserves binary pre-suprema and binary pre-infima.*

Theorem 2. *Let $(P, \leq)$ be a residuated lattice and σ a pre-order on P. Then σ is a pre-congruence on P if and only if*

(i) *$\leq \subseteq \sigma$;*
(ii) *If $a\ \sigma\ b$ and $a'\ \sigma\ b$ then $(a \vee a')\ \sigma\ b$, for all a, b, a' in A;*
(iii) *If $a\ \sigma\ b$ and $a'\ \sigma\ b$ then $(a \wedge a')\ \sigma\ b$, for all a, b, a' in A;*
(iv) *If $a\ \sigma\ b$ and $a'\ \sigma\ b$ then $(a \odot a')\ \sigma\ b$, for all a, b, a' in A;*
(v) *$\odot$ and $\to$ form an adjoint pair, i.e., $c\ \sigma\ (a \to b)$ iff $(a \odot c)\ \sigma\ b$, for all $a, b, c \in A$, where $a \to b =$ pre-suprema$\{c \in P \mid (a \odot c)\ \sigma\ b\}$.*

Proof: Let σ be a pre-congruence on P. Then condition (i) is nothing but condition (1) in Definition 6. Applying Conditions (2) and (4) in Definition 6, we get that $a \vee b$ is a pre-supremum and $a \wedge b$ is a pre-infimum of finite subset of $\{a, b\}$ in P. Hence, and by the definition of pre-suprema and pre-infima, this implies conditions (ii) and (iii). As similar argument, we have for $a \odot b$. Conversely, by condition (i), $a \vee b$ is a pre-upper bound of $A = \{a, b\}$ in (P, σ) and it is a pre-supremum of A by condition (ii), for each binary subset A of P. This implies that Pre-sup$(A) \neq \emptyset$ and also the identity map $id_P : (P, \leq) \longrightarrow (P, \sigma)$ preserves binary pre-suprema. Moreover, condition (i) implies that $a \wedge b$ is a pre-lower bound of A in (P, σ) and it is a pre-infimum of A by condition (ii), for all binary subset A of P. This means that Pre-inf$(A) \neq \emptyset$ and the identity map $id_P : (P, \leq) \longrightarrow (P, \sigma)$ preserves the binary pre-infima of every binary subsets A of P. Also, by condition (iii), it is obvious that id_P preserves $\odot$. Thus we have just proved the conditions (2) and (4) in the definition of a pre-congruence, as required. ∎

A. Directed kernels

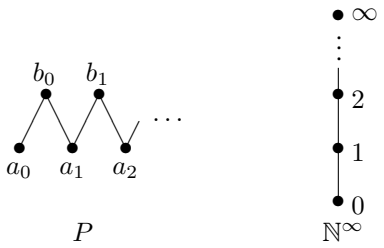
Definition 7. *Let $(P, \leq)$ and $(Q, \sqsubseteq)$ be posets and $f : P \to Q$ be an order-preserving map. The set*

$$\vec{ker} f := \{(a, b) \in P \times P \mid f(a) \sqsubseteq f(b)\},$$

is called the directed kernel or *sub-kernel of f.*

Example 3. *Take the poset $\mathbb{N}^\infty = \mathbb{N} \cup \{\infty\}$ with the natural order, that is, $0 \sqsubseteq 1 \sqsubseteq 2 \sqsubseteq \ldots \sqsubseteq \infty$. Then $\mathbb{N}^\infty$ is a poset. Also, let $(P, \leq)$ be a poset such that $P = \{a_0, a_1, \ldots\} \cup \{b_0, b_1, \ldots\}$ and the order on P is $\leq = \{(a_i, b_i) : \ i \in \mathbb{N}\} \cup \{(a_{i+1}, b_i) : \ i \in \mathbb{N}\} \cup id_P$. Let $f : P \longrightarrow \mathbb{N}^\infty$ be the map*

defined by $f(a_{i+1}) = f(b_i) = i+1$ *for* $i \in \mathbb{N}$, $f(a_0) = 0$. *Trivially,* f *is a poset map. Moreover,* $\vec{\ker} f = \{(a_{i+1}, b_j) \mid j \geq i,\ i,j \in \mathbb{N}\} \cup \{(a_0, b_j) \mid j \in \mathbb{N}\} \cup \{(b_i, a_j) \mid j \geq i,\ j \neq 0,\ i,j \in \mathbb{N}\} \cup \{(a_i, a_j) \mid j \geq i,\ i,j \in \mathbb{N}\} \cup \{(b_i, b_j) \mid j \geq i,\ i,j \in \mathbb{N}\} \cup \mathrm{Id}_P$.



Remark 2. *If* $(P, \leq)$ *and* $(Q, \sqsubseteq)$ *are posets and* $f\colon P \to Q$ *is an order-preserving map, then* $\vec{\ker} f$ *is a pre-order on* P. *In fact, since* $\sqsubseteq$ *is reflexive and transitive,* $\vec{\ker} f$ *is reflexive and transitive as well. Also if* f *is one-one, then* $\vec{\ker} f$ *is a partial order. This is true because for every* $a \in P$, α *(a)* $(a,a) \in \vec{\ker}(\alpha)$, *for every* $a \in P$. $(a,b), (b,b') \in \vec{\ker}\alpha$. *These give* $\alpha(a) \sqsubseteq \alpha(b)$ *and* α *(b) Then* $\alpha(a) \sqsubseteq \alpha(b')$ *and so* (a,b')

Theorem 3. *Let* P *and* Q *be residuated lattices, and* $f\colon P \longrightarrow Q$ *be a residuated lattice homomorphism. Then* $\vec{\ker} f$ *is a pre-congruence.*

Proof: Since f is order-preserving, $\leq \subseteq \vec{\ker} f$. Now, let $a\ \vec{\ker} f\ b$ and $a'\ \vec{\ker} f\ b$ for all $a, b, a' \in P$. This gives that $f(a) \leq f(b)$ and $f(a') \leq f(b)$. So $f(a) \vee f(a') \leq f(b)$. Hence by $f(a \vee a') = f(a) \vee f(a')$, we have $f(a \vee a') \leq f(b)$. Thus $(a \vee a')\ \vec{\ker} f\ b$. Also by $f(a \wedge a') = f(a) \wedge f(a')$, we get $f(a \wedge a') \leq f(b \wedge b')$. This gives that $a \wedge a'\ \vec{\ker} f\ b$. Consequently, by Theorem 2, $\vec{\ker} f$ is a pre-congruence. Also, if $a\ \vec{\ker} f\ (b \to c)$ then $f(a) \leq f(b \to c)$ and so $f(a) \leq f(b) \to f(c)$. By (LR3), one gets $f(a) \odot f(b) \leq f(c)$ then $f(a \odot b) \leq f(c)$, consequently, $a \odot b\ \vec{\ker} f c$. So we get to our result. ∎

In the following we give a characterization of pre-congruences as the directed kernels of residuated lattice homomorphisms.

Theorem 4. *Let* $(P, \leq)$ *be a residuated lattice and* σ *be a pre-order on* P. *Then the following are equivalent:*

1. σ *is a pre-congruence on* P.
2. $(P/\overline{\sigma}, \sigma/\overline{\sigma})$ *is a residuated lattice, the canonical surjection* $\pi\colon (P, \leq) \longrightarrow (P/\overline{\sigma}, \sigma/\overline{\sigma})$ *is a residuated lattice homomorphism and* $\vec{\ker}\pi = \sigma$.
3. *There exist a residuated lattice* $(Q, \preceq)$ *and a residuated lattice homomorphism* $f\colon (P, \leq) \longrightarrow (Q, \preceq)$ *such that* $\vec{\ker} f = \sigma$.

Proof: (1) ⇒ (2): First we prove that $(P/\overline{\sigma}, \sigma/\overline{\sigma})$ has finite joins. To see this, To see this, let $A = \{[a_1], [a_2], \ldots, [a_n]\}$ be a finite subset of $P/\overline{\sigma}$. Then $t := \bigvee \pi_{\overline{\sigma}}^{-1}(A)$ exists in $(P, \leq)$ and so by the hypothesis and the condition (ii) in Theorem 2, t is a pre-supremum of $\pi_{\overline{\sigma}}^{-1}(A)$ in (P, σ). Thus, since $\pi_{\overline{\sigma}}\colon (P, \sigma) \to (P/\overline{\sigma}, \sigma/\overline{\sigma})$ preserves and reflects the relations we obtain that $[t]$ is the supremum of $\pi_{\overline{\sigma}}(\pi_{\overline{\sigma}}^{-1}(A)) = A$ in $(P/\overline{\sigma}, \sigma/\overline{\sigma})$, as required. Next, we prove that $(P/\overline{\sigma}, \sigma/\overline{\sigma})$ has finite meets. Take $s := \bigwedge\{a_1, a_2, \ldots, a_n\}$ in $(P, \leq)$. So by condition (iii) in Theorem 2, s is a pre-infimum of $\{a_1, a_2, \ldots, a_n\}$ in (P, σ). Moreover, it is easy to show that Pre-inf$(\pi_{\overline{\sigma}}^{-1}(A))$ = Pre-inf$(\{a_1, a_2, \ldots, a_n\})$ in (P, σ). This gives that s is a pre-infimum of $\pi_{\overline{\sigma}}^{-1}(A)$ in (P, σ). Now, since $\pi_{\overline{\sigma}}\colon (P, \sigma) \to (P/\overline{\sigma}, \sigma/\overline{\sigma})$ preserves and reflects the relations we obtain that $[s]$ is the infimum of $\pi_{\overline{\sigma}}(\pi_{\overline{\sigma}}^{-1}(A)) = A$ in $(P/\overline{\sigma}, \sigma/\overline{\sigma})$, as required. Since $(P, \leq)$ is a residuated lattice, so it is easy to show that $([a] \odot [b])\sigma/\overline{\sigma}[c] \Leftrightarrow [a]\sigma/\overline{\sigma}([b] \to [c])$ for $[a], [b]$ and $[c]$ in $P/\overline{\sigma}$. It is easy to check that $\pi_{\overline{\sigma}}$ is a homomorphism between residuated lattices. Finally,

$$\begin{aligned}\vec{\ker}\pi_{\overline{\sigma}} &= \{(a,b) \mid \pi_{\overline{\sigma}}(a)\ \sigma/\bar{\sigma}\ \pi_{\overline{\sigma}}(b)\ \} = \{(a,b) \mid [a]\ \sigma/\bar{\sigma}\ [b]\} \\ &= \{(a,b) \mid a\ \sigma\ b\} = \sigma.\end{aligned}$$

(2) ⇒ (3): follows directly from the hypothesis.
(3) ⇒ (1): It is followed immediately from Lemma 3. ∎

B. residuated lattice congruences

In the following, we give a characterization of residuated lattice congruences via a pre-congruence.

Definition 8. *An equivalence relation* R *on a residuated lattice* $(P, \leq)$ *is said to be a residuated lattice congruence if there is an order* $\preceq$ *on* P/R *such that* $(P/R, \preceq)$ *is a residuated lattice and the canonical map* $\pi_R\colon P \to P/R$, $p \mapsto [p]_R$, *is a residuated lattice map.*

Theorem 5. *Let* $(P, \leq)$ *be a residuated lattice and* R *be an equivalence relation on* P. *Then the following are equivalent:*

1. R *is a residuated lattice congruence.*
2. *There exists a pre-congruence* σ *on* P *such that* $R = \sigma \cap \sigma^{-1}$.
3. *There exist a residuated lattice* Q *and a residuated lattice homomorphism* $f\colon P \to Q$ *such that* $R = \ker f$.

Proof: (1) ⇒ (2): Let R be a residuated lattice congruence on P. Then there exists an order $\preceq$ on the quotient set P/R such that $(P/R, \preceq)$ is a residuated lattice and the canonical surjective map $\pi\colon P \to P/R$ is a residuated lattice homomorphism. Take $\sigma = \vec{\ker}\pi$. By Theorem 4, σ is a pre-congruence on P. Moreover, $R = \ker\pi = \vec{\ker}\pi \cap (\vec{\ker}\pi)^{-1} = \bar{\sigma}$.
(2) ⇒ (3): For a pre-congruence σ on P, by Theorem 4, there exist a residuated lattice Q and a surjective residuated lattice homomorphism $f\colon P \to Q$ such that $\sigma = \vec{\ker} f$. Thus $R = \bar{\sigma} = \vec{\ker} f \cap (\vec{\ker} f)^{-1} = \ker f$.
(3) ⇒ (1): By (3) and Theorem 4, $\sigma = \vec{\ker} f$ is a pre-congruence on P. Then again by Theorem 4, $(P/R, \sigma/R)$ is a residuated lattice and the canonical surjection $\pi\colon P \to P/R$ is a residuated lattice homomorphism, where $R = \bar{\sigma} = \ker f$. Thus, R is a residuated lattice congruence. ∎

IV. Conclusions

In this paper, we looked at the quotients of residuated lattices in a different way from that ones already exist. In fact, in set theory we can consider any set with the identity order Δ and then any function between sets is indead an order-preserving map. In this point of view, the kernel of a

function $f : (X, =) \to (Y, =)$ is defined as the set of ordered pairs (a, b) of elements of its domain for which $f(a) = f(b)$. By generalizing this consept to sets with any arbitrary orders, namely posets, we introduced the concept of a directed kernel for an order-preserving map $f : (P, \leq) \to (Q, \sqsubseteq)$ between posets, that is, the set of ordered pairs (a, b) of elements of P with $f(a) \sqsubseteq f(b)$. Having this idea, we introduced the concept of a pre-congruence on a residuated lattice and charecterized them as the intersection of a directed kernel of a residuated lattice homomorphism and its inverse. Using this characterization, we proved that the quotients of residuated lattices are the intersection of such pre-congruences and their inverse.

REFERENCES

[1] P. Hajek, Mathematics of Fuzzy Logic, Dordrecht: Kluwer Academic Publshers, 1998.

[2] A. Iorgulescu, Algebras of logic as BCK algebra, Romania: Academy of Economic Studies Bucharest, 2008.

[3] P. T. Johnstone, Stone Space, Cambridge University Press 1982.

[4] S. Feng and J. Yang, "Congruences in residuated lattices", 7th International Conference on Modelling, Identification and Control, 2015, pp. 18-20.

[5] M. Ward and R. P. Dilworth,"Residuated lattices", Trans. Am. Math. Soc., 45, 1939, pp. 335-354.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)
March 2-4, 2022, Higher Education Complex of Bam

Scattering and Regional Features Fusion Using Collaborative Representation for PolSAR Image Classification

Maryam Imani

Faculty of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran, maryam.imani@modares.ac.ir

***Abstract*— While the collaborative representation has been used for classification of multi-channel images in several works, it is suggested for scattering and spatial features fusion of polarimetric synthetic aperture radar (PolSAR) images in this work. With considering a neighboring region around each pixel of the PolSAR image, its approximation is computed by its adjacent samples in the local region by solving a convex optimization problem. The samples with more similar scattering characteristics will have more important role in the pixel representation. The obtained collaborative representation can be considered as a fused polarimetric-contextual feature space, which can be given to any arbitrary classifier. The experimental results on three real PolSAR images show the good performance of the fused feature space in providing a clear and accurate classification map.**

***Index Terms* — PolSAR, feature fusion, classification, collaborative representation.**

I. INTRODUCTION

There are two types of remote sensors: passive and active. The passive sensors detect electromagnetic radiation reflected from a source such as the sun. In contrast, the active sensors do both tasks of emission and detection of pulses of energy. Synthetic aperture radar (SAR) with capability of imagery in both day and light and any weather conditions is among the main active sensors for various remote sensing applications such surveillance, monitoring, target detection and classification [1]. The polarimetric SAR (PolSAR) sensors can transmit and receive the electromagnetic waves in multiple polarizations of HH, HV, VH and VV. Different land covers show different polarimetric characteristics which can be very useful for discrimination between them. So, PolSAR image classification is a hot and interesting topic in the remote sensing field [2]-[3].

A PolSAR image contains two types of features: the scattering characteristics contained in different polarizations and the spatial characteristics related to relation among the connected pixels. While some works have studied polarimetric feature extraction through applying various target decomposition methods [4]-[5], some other works have studied contextual feature extraction [6]. There is a third group of studies where simultaneously extraction of polarimetric and spatial features or polarimetric-contextual feature fusion is focused [7]-[8].

From the other view, the PolSAR image feature extraction methods are divided into two general groups of manual methods and intelligent ones. In the first group, the mathematical tools or designed operators such as attribute filters [9], Gabor filter banks [10] and morphological profiles [11] are used for manual feature extraction. In the second group, various soft computing tools such as neural networks are utilized to provide an automatic feature extraction. Specially, various deep learning methods such as convolutional neutral networks [12], encoder-decoder [13], residual networks [14] and many different combinations of them [15]-[16] have been introduced for PolSAR image classification.

The main focus in this work is on manual PolSAR feature extraction while the collaborative representation (CR) is suggested for fusion of scattering and contextual feature fusion

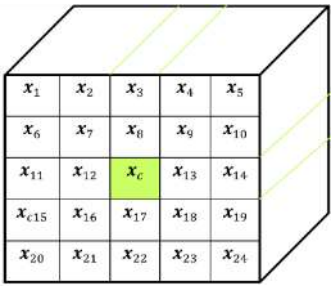


Fig. 1. A central pixel with its 24 neighbors in a 5 × 5 local window.

in local regions. The use of neighborhood information through obtaining a collaborative representation of data has been studied in several works for hyperspectral image analysis [17]-[18] and also in some works for PolSAR image classification [19]-[20]. In most cases, the CR is used to provide a residual term computed from difference of the pixel under test with its approximation calculated by the training samples in a specific classes. But, in this work, CR is used for fusion of polarimetric and spatial features. The obtained CRs computed in neighboring regions can be fed into any arbitrary classifier to provide the classification map. The simple and efficient nearest neighbor classifier is tested in this work. The experiments show the good performance of CR for PolSAR image classification with considering just 1% of datasets as training set.

II. COLLABORATIVE REPRESENTATION BASED FEATURE FUSION

A PolSAR image contains the polarization information of various targets in the ground surface. Each pixel of a PolSAR image can be represented by a 3×3 coherency matrix with $T_{ij};\ i,j = 1,2,3$ elements where $T_{ij} = T_{ji}$. A 9×1 polarimetric feature vector is considered for each PolSAR image as follows:

$$\boldsymbol{x} = \{T_{11}, T_{22}, T_{33}, Re(T_{12}), Re(T_{13}), Re(T_{23}), Im(T_{12}), Im(T_{13}), Im(T_{23})\} \quad (1)$$

where $Re(\cdot)$/ $Im(\cdot)$ denotes the real/imaginary part of $(\cdot)$. In addition to scattering information contained in various polarimetric channels, there is rich spatial information in PolSAR image. Adjacent pixels in neighboring regions have similar scattering characteristics. So, the neighboring pixels belong to the same class with a high probability. To extract this useful information, the use of collaborative representation (CR) is proposed in this work.

Each pixel of the PolSAR cube can be approximated by its neighbors in a local region. Let $\boldsymbol{x}_c$ be a pixel in center of a local window with size of $L \times L$ where there are $p = L^2 - 1$ neighboring pixels. A central pixel in a 5×5 local window with its 24 neighbors is shown in Fig. 1. The central pixel can be approximated by its neighbors $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_p]$ as follows:

$$\boldsymbol{x}_c = \boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{\varepsilon} \quad (2)$$

where $\boldsymbol{\theta}$ is the coefficient vector and $\boldsymbol{\varepsilon}$ denotes the disturbance term. The following optimization problem can be solved to find the coefficient vector:

$$\boldsymbol{\theta} = \arg\min_{\boldsymbol{\theta}^*} \|\boldsymbol{x}_c - \boldsymbol{X}\boldsymbol{\theta}\|_2^2 + \lambda \|\boldsymbol{\Gamma}_{\boldsymbol{x}_c}\boldsymbol{\theta}^*\|_2^2 \quad (3)$$

where the Tikhonov matrix $\boldsymbol{\Gamma}_{\boldsymbol{x}_c}$is calculated by:

$$\boldsymbol{\Gamma}_{\boldsymbol{x}_c} = \begin{bmatrix} \|\boldsymbol{x}_c - \boldsymbol{x}_1\|_2 & & 0 \\ & \ddots & \\ 0 & & \|\boldsymbol{x}_c - \boldsymbol{x}_p\|_2 \end{bmatrix} \quad (4)$$

and λ indicates the regularization parameter. The Tikhonov matrix is used to allow the pixels with more similarity to the central pixel have more role in approximation of it. This is reasonable, because the pixels with close polarimetric signatures likely belong to the same class. The result is:

$$\boldsymbol{\theta} = \left(\boldsymbol{X}^T\boldsymbol{X} + \lambda \boldsymbol{\Gamma}_{\boldsymbol{x}_c}^T \boldsymbol{\Gamma}_{\boldsymbol{x}_c}\right)^{-1} \boldsymbol{X}^T \boldsymbol{x}_c \quad (5)$$

This process is repeated for each pixel of the image to find its associated coefficient vector $\boldsymbol{\theta}$. Then, the CR of each pixel is calculated by $\boldsymbol{x}_c = \boldsymbol{X}\boldsymbol{\theta}$. To compute the CR of PolSAR datasets, $L = 3$ and $\lambda = 0.1$ are used in this work. The use of CR for PolSAR image classification has several advantages such as:

1- Replacing all pixels of the PolSAR image with their CRs is efficacious for speckle noise reduction.
2- Approximation of each pixel with its neighbors in a local region includes the spatial information beside the scattering characteristics. So, CR of a PolSAR data is implicitly a fused feature space containing both polarimetric and spatial features.

The obtained CR can be fed to any arbitrary classifier. The simple and efficient nearest neighbor classifier is used in this work. After providing the classification map, a post processing step can be applied to remove the remained noise pixels. The guided filter as an edge preserving filter is used in this work to smooth the classification map while it preserves the boundaries among different classes. To find more information about this post processing approach, the interested readers are referred to [21]-[22].

III. Experimental results

Three datasets are used to evaluate performance of the PolSAR image features for classification. The Flevoland image acquired by L-band AIRSAR over Flevoland in Netherlands is an agriculture area containing 750×1024 pixels and 15 classes. The SanFrancisco dataset acquired by L-band AIRSAR is a 900 × 1024 image containing four classes. The Oberpfaffenhofen dataset acquired by electronically steered

Table I. Classification results obtained by different feature spaces for the Flevoland dataset.

class			Pol (NP)		CR (NP)		Pol+CR (NP)		Pol (YP)		CR (YP)		Pol+CR (YP)	
No	Name of class	# Total Samples	Acc.	Rel.	Acc.	Rel.	Acc.	Rel.	Acc.	Rel.	Acc.	Rel.	Acc.	Rel.
1	Stembeans	6103	96.23	93.68	96.62	94.55	96.64	94.11	98.03	98.93	97.97	98.83	97.98	98.86
2	Peas	9111	93.52	89.88	95.65	91.37	94.94	90.71	99.44	99.03	99.80	99.45	99.67	99.27
3	Forest	14944	87.55	87.03	89.96	90.00	89.00	89.27	99.97	100	99.99	100	99.98	100
4	Lucerne	9477	92.12	91.41	93.77	92.27	93.14	92.09	94.77	99.97	94.95	99.90	94.77	99.96
5	Wheat	17283	86.45	85.09	87.92	86.28	87.25	86.04	99.64	93.56	99.65	93.73	99.65	93.45
6	Beet	10050	88.43	89.28	89.28	91.25	89.22	90.20	97.59	98.84	97.99	99.23	97.83	99.11
7	Potatoes	15292	83.95	86.56	87.57	88.85	86.59	88.23	99.59	98.61	99.70	98.51	99.63	98.56
8	Bare soil	3078	96.65	98.22	97.04	98.42	96.98	98.29	100	99.90	100	99.90	100	99.81
9	Grass	6269	73.46	81.02	77.97	81.90	76.52	81.46	98.90	95.03	100	95.07	99.86	95.01
10	Rapeseed	12690	78.53	82.61	80.86	85.22	80.28	84.17	90.87	97.46	91.03	97.72	90.60	97.56
11	Barely	7156	92.29	95.46	91.55	97.27	91.91	96.61	99.90	100	99.68	100	99.87	100
12	Wheat 2	10591	83.62	82.05	85.14	84.18	84.89	83.95	99.31	99.96	99.55	99.96	99.51	99.96
13	Wheat 3	21300	93.44	89.48	94.82	91.52	94.34	90.90	99.97	99.40	99.97	99.83	99.97	99.76
14	Water	13476	99.64	99.24	99.65	99.33	99.65	99.31	100	100	100	100	100	100
15	Buildings	476	74.79	100	75.63	100	75.42	100	82.98	100	82.98	100	82.98	100
Average accuracy/ Average reliability			88.04	90.07	89.56	91.50	89.12	91.02	97.40	98.71	97.55	98.81	97.49	98.75
Overall accuracy			88.73		90.40		89.88		98.45		98.58		98.51	
Kappa			87.69		89.51		88.95		98.31		98.45		98.38	

Table II. Z-scores for all pairs of classification methods for the Flevoland dataset.

	Pol (NP)	CR (NP)	Pol+CR (NP)	Pol (YP)	CR (YP)	Pol+CR (YP)
Pol (NP)	0	-25.24	-23.44	-117.81	-119.10	-118.27
CR (NP)	25.24	0	10.41	-105.68	-107.34	-106.41
Pol+CR (NP)	23.44	-10.41	0	-109.84	-111.31	-110.44
Pol (YP)	117.81	105.68	109.84	0	-11.51	-6.20
CR (YP)	119.10	107.34	111.31	11.51	0	8.10
Pol+CR (YP)	118.27	106.41	110.44	6.20	-8.10	0

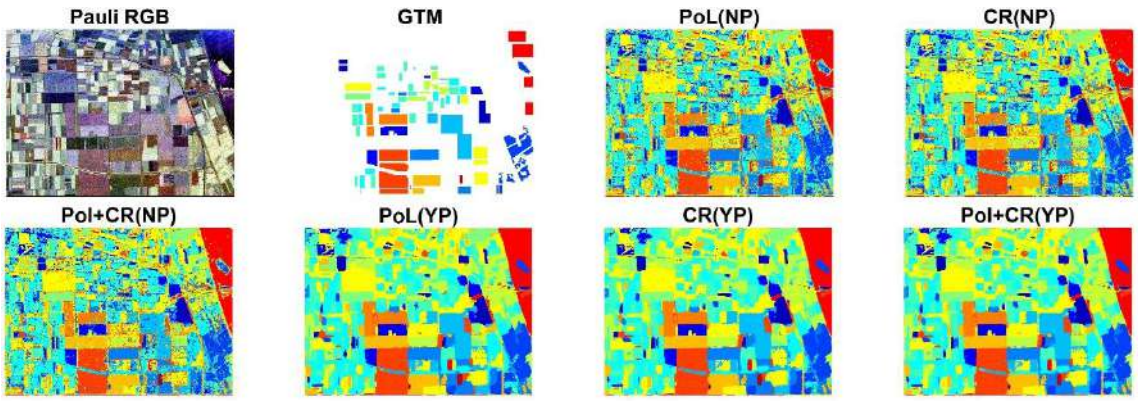


Fig. 2. Classification maps obtained by different feature spaces for the Flevoland dataset.

array radar (ESAR) L-band over Oberpfaffenhofen in Germany has 1297×935 pixels and contains four classes. 1% of data samples is used as training set in all datasets.

The following measures are utilized for assessment of the classification results: accuracy (Acc.) of each class, reliability (Rel.) of each class, average accuracy, average reliability, overall accuracy and kappa coefficient. To show whether the difference between two classification methods is statistically significant or not, the Z score of the McNemars test is computed for all pairs of the methods.

Efficiency of the following features are compared together: Pol: a 9×1 feature vector containing the scatteting channels of the PolSAR cube, CR: the collaborative representation of the PolSAR cube containing 9 fused polarimetric-spatial features, Pol+CR: concatenation of Pol and CR features, which is a 18×1 feature vector containing both polarimetric and fused features. The classification results in two different cases are

Table III. Classification results obtained by different feature spaces for the SanFrancisco dataset.

class			Pol (NP)		CR (NP)		Pol+CR (NP)		Pol (YP)		CR (YP)		Pol+CR (YP)	
No	Name of class	# Total Samples	Acc.	Rel.	Acc.	Rel.	Acc.	Rel.	Acc.	Rel.	Acc.	Rel.	Acc.	Rel.
1	Mountain	61913	77.60	79.96	79.08	82.27	78.67	81.48	88.81	96.63	88.96	96.64	88.94	96.61
2	Grass	135282	61.30	61.79	63.68	63.39	62.74	62.93	51.57	86.58	54.22	86.87	53.12	86.88
3	Sea	348639	98.10	97.71	98.22	97.76	98.19	97.75	99.40	94.69	99.40	94.75	99.39	94.72
4	Building	375766	88.49	88.14	88.75	88.72	88.72	88.49	95.81	86.10	95.76	86.82	95.81	86.53
Average accuracy/ Average reliability			81.38	81.90	82.43	83.03	82.08	82.66	83.90	91.00	84.58	91.27	84.32	91.19
Overall accuracy			87.41		88.00		87.81		90.20		90.58		90.44	
Kappa			81.02		81.93		81.64		84.88		85.48		85.25	

Table IV. Z-scores for all pairs of classification methods for the SanFrancisco dataset.

	Pol (NP)	CR (NP)	Pol+CR (NP)	Pol (YP)	CR (YP)	Pol+CR (YP)
Pol (NP)	0	-20.54	-18.74	-75.69	-86.17	-82.11
CR (NP)	20.54	0	8.28	-60.05	-70.69	-66.56
Pol+CR (NP)	18.74	-8.28	0	-65.09	-75.67	-71.58
Pol (YP)	75.69	60.05	65.09	0	-51.62	-41.30
CR (YP)	86.17	70.69	75.67	51.62	0	29.55
Pol+CR (YP)	82.11	66.56	71.58	41.30	-29.55	0

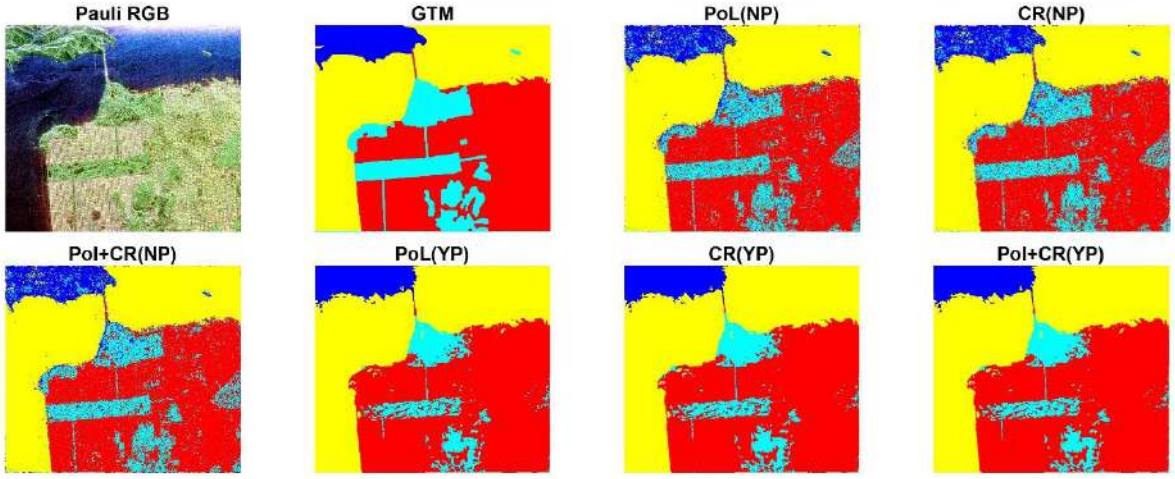


Fig. 3. Classification maps obtained by different feature spaces for the SanFrancisco dataset.

Table V. Classification results obtained by different feature spaces for the Oberpfaffenhofen dataset.

class			Pol (NP)		CR (NP)		Pol+CR (NP)		Pol (YP)		CR (YP)		Pol+CR (YP)	
No	Name of class	# Total Samples	Acc.	Rel.	Acc.	Rel.	Acc.	Rel.	Acc.	Rel.	Acc.	Rel.	Acc.	Rel.
1	Open areas	625029	70.59	70.95	71.41	71.62	71.45	71.45	94.65	67.82	94.65	68.52	94.75	68.18
2	Wood land	202032	49.31	47.83	49.98	48.28	50.01	48.28	78.43	56.90	79.00	58.30	79.04	57.73
3	Built-up areas	190202	26.57	26.43	27.90	27.85	27.39	27.49	16.36	66.25	20.74	68.05	18.76	68.61
4	Road	195432	23.43	23.93	24.33	25.05	23.94	24.78	5.80	75.83	7.33	81.25	6.23	78.46
Average accuracy/ Average reliability			42.48	42.29	43.41	43.20	43.20	43.00	48.81	66.70	50.43	69.03	49.69	68.24
Overall accuracy			52.54		53.43		53.31		65.35		66.38		65.95	
Kappa			27.76		29.07		28.83		40.56		42.59		41.69	

compared: without no post processing (NP) and with applying post processing (YP).

The classification results for Flevoland dataset are represented in Table I. According to the obtained results, generally, applying the post processing improve the classification results

Table VI. Z-scores for all pairs of classification methods for the Oberpfaffenhofen dataset.

	Pol (NP)	CR (NP)	Pol+CR (NP)	Pol (YP)	CR (YP)	Pol+CR (YP)
Pol (NP)	0	-21.50	-24.97	-267.28	-286.76	-278.60
CR (NP)	21.50	0	3.56	-249.56	-269.74	-261.22
Pol+CR (NP)	24.97	-3.56	0	-252.64	-272.59	-264.34
Pol (YP)	267.28	249.56	252.64	0	-90.80	-66.15
CR (YP)	286.76	269.74	272.59	90.80	0	48.43
Pol+CR (YP)	278.60	261.22	264.34	66.15	-48.43	0

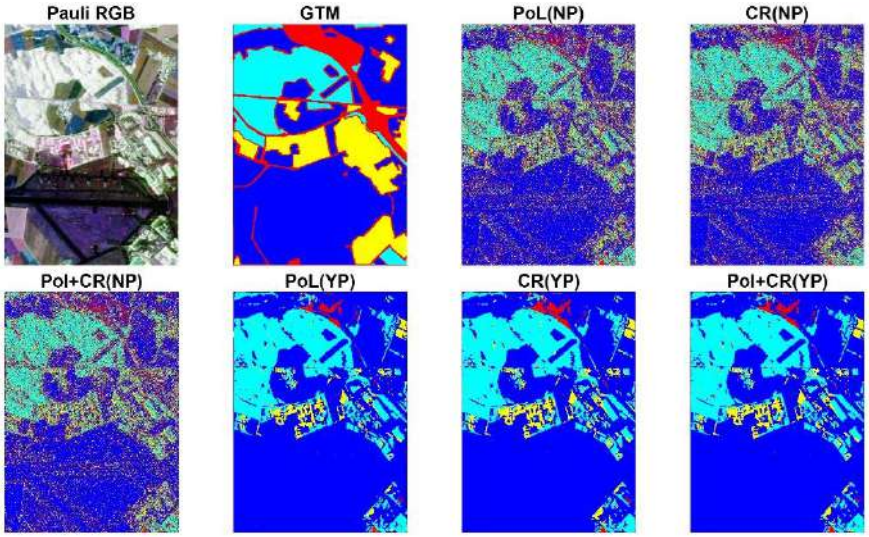


Fig. 4. Classification maps obtained by different feature spaces for the Oberpfaffenhofen dataset.

for all given feature spaces. In other words, all YP cases are better than the NP cases. Among three different feature spaces of Pol, CR, and Pol+CR, the Pol features lead to lowest classification accuracy. This result is expected because considering just the polarimetric features and ignoring the spatial information of neighboring pixels is not appropriate to provide an accurate classification map. The CR method, which fuses both polarimetric and spatial features through relation of neighboring pixels in local regions results in the highest classification accuracy. The Pol+CR, which is the stacked feature space containing both Pol and CR channels has lower classification accuracy compared to CR. The Pol+CR contains both polarimetric and spatial features. But, because the CR feature space itself has both polarimetric and contextual information, adding the polarimetric features to it provides a feature space with higher dimensionality containing redundant information and repeated features, which may degrade the classification result.

The McNemars test results for Flevoland dataset are reported in Table II. As seen, the CR(YP) method provides a positive Z-score much larger than 1.96 with respect to all other methods. It represents that the superiority of CR(YP) with respect to other methods is statistically significant. The Pauli RGB, the ground truth map (GTM) and the obtained classification maps for Flevoland dataset are shown in Fig. 2. While the NP cases show high noise pixels, the YP cases are much cleaner than the NP ones. Among Pol(NP), CR(NP) and Pol+CR(NP), the classification map of CR(NP) contains less incorrect labeled pixels. Similarly, among Pol(YP), CR(YP) and Pol+CR(YP), the classification map of the CR(YP) method is the most accurate.

The classification results for SanFrancisco dataset are reported in Tables III-IV and Fig. 3, and the classification results for Oberpfaffenhofen dataset are represented in Tables V-VI and Fig. 4. The results for these two datasets are similar to previous dataset. As a general conclusion, the use of the CR feature space, which contains the fused polarimetric and regional features together with applying the post processing filters leads to best classification results for all datasets.

IV. CONCLUSION

The polarimetric features and regional characteristics of the neighboring pixels are fused through the collaborative representation (CR) in this work. According to the done experiments, CR can be operated as a rich polarimetric-spatial feature space, which can fed into a simple classifier to find an accurate classification map. The speckle noise is reduced in CR. In addition, CR contains the contextual information computed from the neighboring pixels in the local regions. So, it can be a good candidate for PolSAR image classification.

REFERENCES

[1] M. Limbach, B. Gabler, R. Horn, A. Kosc, A. Di Maria and R. Scheiber, "P-band antenna array for airborne SAR application and DBF SAR demonstration," *2015 9th European Conference on Antennas and Propagation (EuCAP)*, pp. 1-5, 2015.

[2] R. Gui, X. Xu, L. Wang, R. Yang and F. Pu, "Eigenvalue Statistical Components-Based PU-Learning for PolSAR Built-Up Areas Extraction and Cross-Domain Analysis," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 3192-3203, 2020.

[3] D. A. Yocky, R. D. West and R. E. Abbott, "Comparison of PolSAR Surface Measurements From Underground Chemical Explosions to Recorded and Predicted Surface Ground Motion," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 165-174, 2021.

[4] N. Besic, G. Vasile, J. Chanussot and S. Stankovic, "Polarimetric Incoherent Target Decomposition by Means of Independent Component Analysis," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 3, pp. 1236-1247, March 2015.

[5] B. Zou, D. Lu, L. Zhang and W. M. Moon, "Eigen-Decomposition-Based Four-Component Decomposition for PolSAR Data," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 3, pp. 1286-1296, March 2016.

[6] L. Zhang, L. Sun and W. M. Moon, "Polarimetric SAR image classification based on contextual sparse representation," *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2015, pp. 1837-1840.

[7] M. Imani, "Polarimetric SAR Classification Using Ridge Regression-Based Polarimetric-Spatial Feature Extraction,"26th International Computer Conference, Computer Society of Iran, Tehran, Iran, 3-4 March 2021.

[8] J. Shi, H. Jin and X. Li, "A Novel Multi-Feature Joint Learning Method for Fast Polarimetric SAR Terrain Classification," in *IEEE Access*, vol. 8, pp. 30491-30503, 2020.

[9] A. Tombak, İ. Türkmenli, E. Aptoula and K. Kayabol, "Pixel-Based Classification of SAR Images Using Feature Attribute Profiles," in *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 4, pp. 564-567, April 2019.

[10] L. Zhang, Y. Huang, L. Ferro-Famil and W. Wu, "Effect of polarimetric information on time-frequency analysis using spaceborne SAR image," *EUSAR 2021; 13th European Conference on Synthetic Aperture Radar*, pp. 1-6, 2021.

[11] P. R. Marpu, K. Chen, C. Chu and J. A. Benediktsson, "Spectral-spatial classification of polarimetric SAR data using morphological profiles," *2011 3rd International Asia-Pacific Conference on Synthetic Aperture Radar (APSAR)*, pp. 1-3, 2011.

[12] D. Song *et al*., "A Novel Marine Oil Spillage Identification Scheme Based on Convolution Neural Network Feature Extraction From Fully Polarimetric SAR Imagery," in *IEEE Access*, vol. 8, pp. 59801-59820, 2020.

[13] F. Gu, H. Zhang, C. Wang and B. Zhang, "Residual encoder-decoder network introduced for multisource SAR image despeckling," *2017 SAR in Big Data Era: Models, Methods and Applications (BIGSARDATA)*, pp. 1-5, 2017.

[14] C. Wang, J. Pei, X. Liu, Y. Huang and J. Yang, "A deep deformable residual learning network for SAR image segmentation," *2021 IEEE Radar Conference (RadarConf21)*, pp. 1-5, 2021.

[15] M. Imani, "Integration of the *k*-nearest neighbours and patch-based features for PolSAR image classification by using a two-branch residual network," Remote Sensing Letters, vol. 12, no. 11, pp. 1112–1122, 2021.

[16] R. Qin, X. Fu and P. Lang, "PolSAR Image Classification Based on Low-Frequency and Contour Subbands-Driven Polarimetric SENet," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 4760-4773, 2020.

[17] M. Imani, "Sparse and collaborative representation-based anomaly detection," Signal Image and Video Processing, vol. 14, no. 8, pp. 1573–1581, 2020.

[18] M. Imani and H. Ghassemian, "Weighted Joint Collaborative Representation Based On Median-Mean Line and Angular Separation," IEEE Transactions on Geoscience and Remote Sensing, vol. 55, no. 10, pp. 5612 – 5624, Oct. 2017.

[19] J. Geng, J. Fan, H. Wang, A. Fu and Y. Hu, "Joint collaborative representation for polarimetric SAR image classification," *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 3066-3069, 2016.

[20] J. Geng, H. Wang, J. Fan, X. Ma, B. Wang, "Wishart distance-based joint collaborative representation for polarimetric SAR image classification," IET Radar Sonar Navig., vol. 11, pp. 1620-1628, 2017.

[21] X. Kang, S. Li and J. A. Benediktsson, "Spectral–Spatial Hyperspectral Image Classification With Edge-Preserving Filtering," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 5, pp. 2666-2677, May 2014.

[22] M. Imani, "A Random Patches Based Edge Preserving Network for Land Cover Classification Using Polarimetric Synthetic Aperture Radar Images," International Journal of Remote Sensing, vol. 42, no. 13, pp. 4946–4964, 2021.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Soaker ideals in MV-algebras

1st Fereshteh Forouzesh, 2nd Naser Hosseini

1st Department of Pure Mathematics and Calculations , Faculty of Mathematics, Higher Education Complex of Bam, Kerman, Iran, frouzesh@bam.ac.ir

2nd Department of Pure Mathematics, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman, Kerman, Iran, nhoseini@uk.ac.ir

***Abstract*— In this paper, we define soaker ideals in an MV-algebra, and study the relationships between soaker ideals and the other ideals in an involutive MV-algebras.**

***Index Terms*— finitely meet-irreducible, involutive, (soaker, join-soaker) ideal.**

I. Introduction and Preliminaries

MV-algebras are introduced by C. C. Chang in 1958 [1] to give an algebraic counterpart of the multiple-valued Łukasiewicz propositional logic.

After their introduction by Chang, MV-algebras free themselves from the bonds of logic and become an autonomous mathematical discipline with deep connections to several other branches of mathematics. For example, in 1986 D. Mundici proved that the category of lattice ordered abelian groups with strong unit is categorical equivalent to the category of MV-algebras ([9]). This result is very important because lattice ordered abelian groups don't set up an equational variety unlike MV-algebras. In this way more complicated properties in groups language can became simpler in MV-algebras language. Moreover, the study of normal forms for Lukasiewicz logic brought to a deep relation between MV-algebras and toric varieties through the concept of Schauder bases, which are the affine versions of a complex of nonsingular cones.
In this paper, we recall some facts concerning MV-algebras, introduce soaker ideals in an MV-algebra A, investigate some relationships between the soaker ideals and the other ideals of an involutive MV-algebra and give characterization of soaker ideals and study some properties of them.
In all the sections, we recollect some definitions and results which will be used in the following:

Definition 1: [1] An MV-algebra is a structure (A, $\oplus$, *, 0), where $\oplus$ is a binary operation, *, is a unary operation, and 0 is a constant such that the following axioms are satisfied for any $a, b \in A$:
($MV1$) (A, $\oplus$, 0) is an abelian monoid,
($MV2$) $(a^*)^* = a$,
($MV3$) $0^* \oplus a = 0^*$,
($MV4$) $(a^* \oplus b)^* \oplus b = (b^* \oplus a)^* \oplus a$.

Note that $1 = 0^*$ and the auxiliary operation $\odot$ is defined as follow:

$$x \odot y = (x^* \oplus y^*)^*.$$

Lemma 1: [10] Let A be an MV-algebra. For $x, y \in A$, the following conditions are equivalent:
(1) $x^* \oplus y = 1$,
(2) $x \odot y^* = 0$,
(3) There is an element $z \in M$ such that $x \oplus z = y$,
(4) $y = x \oplus (y \ominus x)$.
For any two elements $x, y \in A$, $x \leq y$ iff x and y satisfy the equivalent conditions (1)-(4) in the above lemma, which $\leq$ is natural order of A.

We recall that the natural order determines a bounded distributive lattice structure such that

$$x \vee y = x \oplus (x^* \odot y) = y \oplus (x \odot y^*) \quad \text{and} \quad x \wedge y = x \odot (x^* \oplus y) = y \odot (y^* \oplus x)$$

Lemma 2: [10] In each MV-algebra A, the following relations hold for all $x, y, z \in A$:
(1) $x \leq y$ if and only if $y^* \leq x^*$,
(2) If $x \leq y$, then $x \oplus z \leq y \oplus z$ and $x \odot z \leq y \odot z, x \wedge z \leqslant y \wedge z$,
(3) $x \leq y$ if and only if $x^* \oplus y = 1$ if and if $x \odot y^* = 0$,
(4) $x, y \leq x \oplus y$ and $x \odot y \leq x, y$, $x \leq nx = x \oplus x \oplus \cdots \oplus x$ and $x^n = x \odot x \odot \cdots \odot x \leq x$,
(5) $x \oplus x^* = 1$ and $x \odot x^* = 0$,
(6) If $x \leq y$ and $z \leq t$, then $x \oplus z \leq y \oplus t$,
(7) $x \odot (y \vee z) = (x \odot y) \vee (x \odot z)$,
(8)$x \wedge (y \oplus z) \leqslant (x \wedge y) \oplus (x \wedge z)$, *hence* $x \wedge (x_1 \oplus \cdots \oplus x_n) \leqslant (x \wedge x_1) \oplus \cdots \oplus (x \wedge x_n)$, *for all* $x_1, \ldots, x_n \in A$.

Definition 2: [2] An ideal of an MV-algebra A is a nonempty subset I of A satisfying the following conditions:
($I1$) If $x \in I$, $y \in A$ and $y \leq x$ then $y \in I$,
($I2$) If $x, y \in I$, then $x \oplus y \in I$.
We denote by $Id(A)$ the set of ideals of an MV-algebra A.

Definition 3: [2] Let P be an ideal of an MV-algebra A. Then P is a proper ideal of A if $P \neq A$. An ideal I of an MV-algebra A is called

• [2] prime ideal if for all $x, y \in A, x \wedge y \in P$, then $x \in P$ or $y \in P$.
We denote by $Spec(A)$ the set of prime ideals of an MV-algebra A.
• [10] finitely meet-irreducible if $I \cap J \subseteq P$, then $I \subseteq P$ or $J \subseteq P$, for all $I, J \in Id(A)$.

Lemma 3: [10] In an MV-algebra A, ideal P is finitely meet-irreducible if and only if P is prime ideal of A.

Definition 4: [2], [8] Let X and Y be two MV-algebras. A function $f : X \to Y$ is called homomorphism of MV-algebras if and only if
(1) $f(0) = 0$,
(2) $f(x \oplus y) = f(x) \oplus f(y)$,
(3) $f(x^*) = (f(x))^*$.

Remark 1: [2] We recall that for a nonempty subset $X \subseteq A$, the smallest ideal of A which contains X, i.e., $\bigcap\{I \in Id(A) : X \subseteq I\}$, is said to be the ideal of A generated by X and will be denoted by $(X]$.

Remark 2: [10] Let $X \subseteq A$. We have
$(1)(X] = \{a \in A : a \leqslant x_1 \oplus x_2 \oplus ... \oplus x_n,$ *for some* $n \in \mathbb{N}$ *and* $x_1, ..., x_n \in X\}$.
$I \in Id(A)$ is called a finitely generated ideal, if $I = (x_1, ..., x_n]$, *for some*
$x_1, x_2, ..., x_n \in A$ *and* $n \in \mathbb{N}$.
In particular, $(a] = \{x \in A : x \leqslant na,$ *for some* $n \in \mathbb{N}\}$.
(2) *For* $I_1, I_2 \in Id(A), I_1 \vee I_2 = \{x \in A : \exists a_i \in I; x \leqslant a_1 \oplus a_2\}$.
(3) $(a] \wedge (b] = (a \wedge b]$.
(4) if $a \leq b$, then $(a] \subseteq (b]$.

Proposition 1: [10] Let $f : A \to B$ be a homomorphism of MV-algebras.
(i) if J is a proper ideal of B, then $f^{-1}(J)$ is a proper ideal of A.
(ii) if f is onto and I is an ideal of A, then $f(I)$ is an ideal of B.

Lemma 4: [10] Let $f : A \to B$ be an onto homomorphism of MV-algebras A, B and $\{I_i\}_{i \in I}$ be a family of ideals of A. Then we have $f(\vee_{i \in I} I_i) = \vee_{i \in I} f(I_i)$.

Definition 5: [2] Let X be a nonempty subset of MV-algebra A and $Ann_A(X)$ be the annihilator of X defined by

$$Ann_A(X) = \{a \in A : a \wedge x = 0, \forall x \in X\}$$

II. Soaker ideals in MV-algebras

In the sequel, A is an MV-algebras.

Definition 6: Let I be an ideal of A. I is called a soaker ideal of A, if
(i) $I \neq \{0\}$
(ii) For $\{I_i\}_{i \in J} \subseteq Id(A)$, If $I \subseteq \bigvee_{i \in J} I_i$, then $I \subseteq I_i$, for some $i \in J$.

Example 1: Let $A = \{0, a, b, 1\}$, where $0 < a, b < 1$. Define $\odot$, $\oplus$ and $*$ as follows:

$\odot$	0	a	b	1
0	0	0	0	0
a	0	a	0	a
b	0	0	b	b
1	0	a	b	1

$\oplus$	0	a	b	1
0	0	a	b	1
a	a	a	1	1
b	b	1	b	1
1	1	1	1	1

$*$	0	a	b	1
	1	b	a	0

Then $(A, \oplus, \odot, *, 0, 1)$ is an MV-algebra [8], it is clear that $I_1 = \{0\}$, $I_2 = \{0, a\}$, $I_3 = \{0, b\}$ and $I_4 = A$ are ideals of A and I_2 and I_3 are soaker ideals, while A is not a soaker ideal because $A \subseteq I_2 \vee I_3$. But $A \nsubseteq I_2$ and $A \nsubseteq I_3$.

Example 2: Consider $S_1 = \{0, 1\}$, $S_2 = \{0, 1/2, 1\}$. Then $A = S_1 \times S_2$ with operations $(x, y) \oplus (z, t) = (min\{1, x + z\}, min\{1, y + t\})$ and $(x, y)^* = (1 - x, 1 - y)$ is an MV-algebra. We have $I_0 = \{(0, 0)\}$, $I_1 = \{(0, 0), (1, 0)\}$, $I_2 = \{(0, 0), (0, 1/2), (0, 1)\}$ and A are ideals of A. It can be easily verified that I_1, I_2 are soaker ideals, while A is not a soaker ideal.

Example 3: Let $A = \{0, a, b, c, d, e, f, 1\}$ is rectangular cube such that $0 < a, d < e < 1$, $0 < a, b < c < 1$, $0 < b, d < f < 1$, $0 < d < e, f < 1$, $0 < a < c, e < 1$ and $0 < b < c, f < 1$.
Define $\oplus$ and $*$ as follows:

$\oplus$	0	a	b	c	d	e	f	1
0	0	a	b	c	d	e	f	1
a	a	a	c	c	e	e	1	1
b	b	c	b	c	f	1	f	1
c	c	c	c	c	1	1	1	1
d	d	e	f	1	d	e	f	1
e	e	e	1	1	e	e	1	1
f	f	1	f	1	f	1	f	1
1	1	1	1	1	1	1	1	1

$*$	0	a	b	c	d	e	f	1
	1	f	e	d	c	b	a	0

Then $(A, \oplus, *, 0, 1)$ is an MV-algebra [8], it is clear that $I_0 = \{0\}$, $I_1 = \{0, a\}$, $I_2 = \{0, d\}$, $I_3 = \{0, b\}$, $I_4 = \{0, a, d, e\}$, $I_5 = \{0, a, b, c\}$, $I_6 = \{0, b, d, f\}$ and $I_7 = A$ are ideals of A. We can easily see that I_1, I_2 and I_3 are soaker ideals of A,but I_4, I_5, I_6 and I_7 are not soaker ideals. For example, $I_4 \subseteq I_1 \vee I_2$. But $I_4 \nsubseteq I_1$ and $I_4 \nsubseteq I_2$. Hence I_4 is not soaker ideal of A.

Example 4: Let $G = \oplus\{Z_i / i \in \mathbb{N}\}$ be the lexicographic product of denumerable infinite copies of the abelian l-group $\mathbb{Z}$ of the relative integers and $e^i \in G$ such that $e^i_k = 0$ if $k \neq i$ and $e^i_k = 1$ if $k = i$.
Consider the perfect MV-algebra $A = \Gamma(G)$, where Γ is a functor from the category of abelian l-groups to the category

of prefect MV-algebras [5].
If we set $P_i =< (0, e^i) >$, then $P_i \subseteq P_j$, for $i > j$, hence the set of all soaker ideals of A is $\{P_i / i \in \mathbb{N}\}$.

Theorem 1: S is a soaker ideal of A if and only if $S = \bigvee_{i\in J} I_i$ implies $S = I_i$, for some $i \in J$.

Proof: Let S be a soaker ideal and $S = \bigvee_{i\in J} I_i$. Since $S \subseteq \bigvee_{i\in I} I_i$, we get $S \subseteq I_i \subseteq \bigvee_{i\in J} I_i = S$. Hence $S = I_i$.
Conversely, suppose that $S \subseteq \bigvee_{i\in J} I_i$. Then $S \cap \bigvee_{i\in J} I_i = \bigvee_{i\in J}(S \cap I_i) = S$, by hypothesis, $S \cap I_i = S$, for some $i \in J$. Thus $S \subseteq I_i$, for some $i \in J$. Therefore S is a soaker ideal of A. ∎

Lemma 5: In an MV-algebra A, we have $(a] \vee (b] = (a \vee b] = (a \oplus b]$.

Proof: Since $a, b \leq a \vee b \leq a \oplus b$, by Remark 2 (4), we deduce that $(a], (b] \subseteq (a \vee b] \subseteq (a \oplus b]$. Hence $(a] \vee (b] \subseteq (a \vee b] \subseteq (a \oplus b]$.
Conversely, let $x \in (a \oplus b]$. It follows from Remark 2 $(1), (2)$ that for some $n \geq 1$,

$$x \leq n(a \oplus b) = na \oplus nb \in (a] \vee (b].$$

Hence $x \in (a] \vee (b]$. That is, $(a \oplus b] \subseteq (a] \vee (b]$. Thus we have $(a] \vee (b] = (a \oplus b]$. So $(a] \vee (b] = (a \vee b] = (a \oplus b]$. ∎
Since for any ideal I, $I = \bigvee_{a\in I}(a]$, we have:

Corollary 1: Every soaker ideal is principal.

Setting $C = \{z \in A : \forall x, y \in A, z \leq x \oplus y$ implies $\exists m \in \mathbb{N}, z \leq mx$ or $\exists n \in \mathbb{N}, z \leq ny\}$, we have:

Theorem 2: $(a]$ is soaker if and only if $a \in C$.

Proof: Suppose that $(a]$ is soaker. If $a \leq x \oplus y$, then by Remark 2 (4), (3), we get $(a] \subseteq (x] \vee (y]$. Hence $(a] \subseteq (x]$ or $(a] \subseteq (y]$. So $a \leq mx$, for some $m \in \mathbb{N}$ or $a \leq ny$, for some $n \in \mathbb{N}$. Thus $a \in C$.

Now, suppose $a \in C$. If $(a] \subseteq \vee_{\alpha\in I}(a_\alpha]$, then $a \in \vee_{\alpha\in I}(a_\alpha]$. We obtain $a \leq a_{\alpha_1} \oplus \cdots \oplus a_{\alpha_k}$. By hypothesis with some manupulations one obtains, $a \leq m_1 a_{\alpha_1}$ or $a \leq m_2 a_{\alpha_2}$ or $\cdots$ or $a \leq m_k a_{\alpha_k}$. We conclude that $(a] \subseteq (a_{\alpha_1}]$ or $(a] \subseteq (a_{\alpha_2}]$ or $\cdots$ or $(a] \subseteq (a_{\alpha_k}]$. Thus $(a]$ is soaker. ∎

Lemma 6: If $a \in C$, then for all $n \in \mathbb{N}$, $na \in C$; and if there exists $n \in \mathbb{N}$ such that $na \in C$, then $a \in C$

Proof: Suppose that $a \in C$. Let $n \in \mathbb{N}$ and $na \leq x \oplus y$. Since $a \leq na$, we get $a \leq x \oplus y$. So $a \leq kx$ or $a \leq ly$. Thus $na \leq knx$ or $na \leq lny$. Therefore $na \in C$.

Now, suppose there exists $n \in \mathbb{N}$ such that $na \in C$. If $a \leq x \oplus y$, then $na \leq nx \oplus ny$. By hypothesis, we have $na \leq (nk)x$ or $na \leq (nl)y$. Since $a \leq na$, we deduce that $a \leq (nk)x$ or $a \leq (nl)y$. Therefore $a \in C$. ∎

Theorem 3: Let $f : A \to B$ be a homomorphism of MV-algebras and S be a soaker ideal of B. Then $f^{-1}(S)$ is a soaker ideal of A.

Proof: Clearly, $f^{-1}(S)$ is an ideal of A. If $f^{-1}(S) = \{0\}$, then $f(\{0\}) = S$, and so $\{0\} = S$, which is a contradiction. Hence $f^{-1}(S) \neq \{0\}$.
Let $\{I_i\}_{i\in J}$ be a family of ideals of A such that $f^{-1}(S) \subseteq \vee_{i\in J} I_i$. It follows from Lemma 4 that $S \subseteq f(\vee_{i\in J} I_i) = \vee_{i\in J} f(I_i)$. Now, since S is a soaker ideal of B, we get $S \subseteq f(I_i)$, for some $i \in J$. Hence $f^{-1}(S) \subseteq I_i$, for some $i \in J$. Thus $f^{-1}(S)$ is a soaker ideal of A. ∎

Theorem 4: Let S be a soaker ideal of A and $f : A \to B$ be an isomorphism of MV-algebras. Then $f(S)$ is a soaker ideal of B.

Proof: Let $\{I_i\}_{i\in J}$ be a family of ideals of B such that $f(S) \subseteq \vee_{i\in J} I_i$. Since f is onto, we get $I_i = f(I_i')$, for some $I_i' \in Id(A)$, and so $f(S) \subseteq \vee_{i\in I} f(I_i')$, for some $I_i' \in Id(A)$. By Lemma 4, we have $f(S) \subseteq f(\vee_{i\in J} I_i')$ and since f is one to one, so $S \subseteq \vee_{i\in J} Ii'$. By hypothesis, $S \subseteq I_i'$, for some $i \in J$. Hence $f(S) \subseteq f(I_i') = I_i$, for some $i \in J$. Thus, $f(S)$ is a soaker ideal of B. ∎

Definition 7: An MV-algebra A is called involutive if $I = Ann(Ann(I))$, for every ideal I of A.

Example 5: (i) MV-algebra in Example 1 is involutive.
(ii) MV-algebra in Example 4 is involutive.

Example 6: Let $C = \{0, c, 2c, 3c, \ldots, 1 - 2c, 1 - c, 1\}$ be the MV-algebra defined in [1] with operations as follows:
if $x = nc$ and $y = mc$, then $x \oplus y := (m + n)c$,
if $x = 1 - nc$ and $y = 1 - mc$, then $x \oplus y := 1$,
if $x = nc$ and $y = 1 - mc$ and $m \leq n$, then $x \oplus y := 1$,
if $x = nc$ and $y = 1 - mc$ and $n < m$, then $x \oplus y := 1 - (m - n)c$,
if $x = 1 - mc$ and $y = nc$ and $m \leq n$, then $x \oplus y := 1$,
if $x = 1 - mc$ and $y = nc$ and $n < m$, then $x \oplus y := 1 - (m - n)c$,
if $x = nc$, then $x^* := 1 - nc$,
if $x = 1 - nc$, then $x^* := nc$.

We see that C is a prefect MV-algebra (i.e, C has only a maximal ideal) and it has only three ideals: $\{0\}$, $M = \{0, c, 2c, 3c, \ldots\}$ and C. Since $Ann(Ann(M)) = Ann(\{0\}) = A \neq M$, C is not involutive.

Lemma 7: Let A be an MV-algebra and $I, J \in Id(A)$. We have
(i) $I \subseteq Ann(Ann(I))$,
(ii) if $I \subseteq J$, then $Ann(J) \subseteq Ann(I)$,
(iii) $Ann(I \cap J) = Ann(I) \cap Ann(J)$,
(iv) $Ann(I \vee J) \supseteq Ann(I) \cap Ann(J)$,

Proof: (i) Let $a \in I$. Suppose that $x \in Ann(I)$. Hence $x \wedge t = 0$, for all $t \in I$. Since $a \in I$, it follows that $x \wedge a = 0$, for $x \in Ann(I)$. Thus $a \in Ann(Ann(I))$. Therefore we get $I \subseteq Ann(Ann(I))$.
(ii) Suppose that $a \in Ann(J)$. Hence $a \wedge x = 0$, for all $x \in J$. We get $a \wedge x = 0$, for all $x \in I$. Then we obtain $a \in Ann(I)$. Thus $Ann(J) \subseteq Ann(I)$.
(iii) Suppose that $a \in Ann(I \cap J)$ if and only if $a \wedge x = 0$, for all $x \in I \cap J$ if and only if $a \wedge x = 0$, for all $x \in I$ and $x \in J$ if and only if $a \in Ann(I)$ and $a \in Ann(J)$ if and only if $a \in Ann(I) \cap Ann(J)$.
(iv) Suppose that $a \in Ann(I) \cap Ann(J)$. Hence $a \wedge t = 0$ and $a \wedge s = 0$, for all $t \in I$ and $s \in J$.
Now, for all $x \in I \vee J$, by Remark 2 (2), we get $x \leq c \oplus b$, for some $c \in I$ and $b \in J$. It follows from Lemma 2 (2), (8) that

$$a \wedge x \leq a \wedge (c \oplus b) \leq (a \wedge c) \oplus (a \wedge b) = 0$$

Hence $a \in Ann(I \vee J)$. Thus $Ann(I \vee J) \supseteq Ann(I) \cap Ann(J)$. ∎

Lemma 8: Let A be involutive. We have $Ann(I) \vee Ann(J) \subseteq Ann(I \cap J)$.

Proof: Since A is an involutive, by Lemma 7 (iii), (iv), we have

$$\begin{aligned} I \cap J &= Ann(Ann(I \cap J), \\ &= Ann(Ann(I) \cap Ann(J)), \\ &= Ann(Ann(I)) \cap Ann(Ann(J)), \\ &\subseteq Ann(Ann(I) \vee Ann(J)). \end{aligned}$$

Hence $Ann(Ann(I \cap J)) \subseteq Ann(Ann(I) \vee Ann(J))$, so by Lemma 7 (ii), we get $Ann(Ann(Ann(I) \vee Ann(J))) \subseteq Ann(Ann(Ann(I \cap J)))$. Thus we obtain $Ann(I) \vee Ann(J) \subseteq Ann(I \cap J)$. ∎

Theorem 5: Let A be an involutive MV-algebra.
Every non-zero proper ideal of A is finitely meet-irreducible ideal if and only if it is a soaker ideal of A.

Proof: Let P is soaker ideal of A and $I \cap J \subseteq P$, where $I, J \in Id(A)$. It follows from Lemma 7 (iii) and (iv) that

$$\begin{aligned} Ann(P) &\subseteq Ann(I \cap J), \\ &= Ann(I) \cap Ann(J), \\ &\subseteq Ann(I) \vee Ann(J). \end{aligned}$$

Since $Ann(P)$ is soaker ideal, $Ann(P) \subseteq Ann(I)$ or $Ann(P) \subseteq Ann(J)$. Then by Lemma 7 (ii), we get $Ann(Ann(P)) \supseteq Ann(Ann(I))$ or $Ann(Ann(P)) \supseteq Ann(Ann(J))$. Then we obtain $P \supseteq I$ or $P \supseteq J$.
Conversely, suppose that P is meet-irreducible ideal of A. Let $P \subseteq I_1 \vee I_2 \vee \cdots \vee I_k$, where $I_i \in Id(A)$, for $i = 1, \cdots k$. It follows from Lemma 7 (ii) and (iv) that

$$Ann(P) \supseteq Ann(I_1 \vee \cdots \vee I_k) \supseteq Ann(I_1) \cap \cdots \cap Ann(I_k).$$

Since $Ann(P)$ is an ideal, so by hypothesis, $Ann(P) \supseteq Ann(I_1)$ or $Ann(P) \supseteq Ann(I_2)$,..., $Ann(P) \supseteq Ann(I_k)$. Now, by Lemma 7 (ii), we get $Ann(Ann(P)) \subseteq Ann(Ann(I_1))$ or ....or $Ann(Ann(P)) \subseteq Ann(Ann(I_k))$. We obtain $P \subseteq I_1$ or ... or $P \subseteq I_k$. ∎

Theorem 6: If ideal P is soaker, then $Ann(P)$ is finitely meet-irreducible.

Proof: Suppose P is soaker ideal and $I \cap J \subseteq Ann(P)$. It follows from Lemma 7 (ii) that $Ann(Ann(P)) \subseteq Ann(I \cap J)$. By Lemma 7 (i), (iii) and (iv), we have

$$\begin{aligned} P &\subseteq Ann(Ann(P)), \\ &\subseteq Ann(I \cap J), \\ &= Ann(I) \cap Ann(J), \\ &\subseteq Ann(I) \vee Ann(J). \end{aligned}$$

Since P is soaker ideal, $P \subseteq Ann(I)$ or $P \subseteq Ann(J)$. By Lemma 7 (ii), we have $Ann(Ann(I)) \subseteq Ann(P)$ or $Ann(Ann(J)) \subseteq Ann(P)$. It follows from Lemma 7 (i) that $I \subseteq Ann(P)$ or $J \subseteq Ann(P)$. ∎

Theorem 7: If A is involutive and $Ann(P)$ is finitely meet-irreducible, then ideal P is soaker.

Proof: Let P is finitely meet-irreducible and $P \subseteq I_1 \vee I_2 \vee \cdots \vee I_k$, where $I_i \in Id(A)$, for $i = 1, \cdots k$. Hence by Lemma 7 (ii) and (iv), we get

$$Ann(P) \supseteq Ann(I_1 \vee \cdots \vee I_k) \supseteq Ann(I_1) \cap \cdots \cap Ann(I_k).$$

Since $Ann(P)$ is finitely meet-irreducible, $Ann(P) \supseteq Ann(I_1)$ or $Ann(P) \supseteq Ann(I_2)$,..., $Ann(P) \supseteq Ann(I_k)$. Now, by Lemma 7 (ii), we get $Ann(Ann(P)) \subseteq Ann(Ann(I_1))$ or ....or $Ann(Ann(P)) \subseteq Ann(Ann(I_k))$. We obtain $P \subseteq I_1$ or ... or $P \subseteq I_k$. ∎
By Theorem 5 and Lemma 3, we obtain the following corollary:

Corollary 2: Let A be an involutive MV-algebra. P is soaker ideal if and only if P is a prime ideal.

CONCLUSION

In this paper, we introduced soaker ideals in MV-algebras and investigated some properties of them. We stated the relationships between soaker ideals in a involutive MV-algebras.

REFERENCES

[1] Chang, C. C., " Algebraic analysis of many valued logic", Trans. Amer. Math. Soc. 88 (1958), 467-490.
[2] Cignoli, R. D'Ottaviano, I. M. L.Mundici, D. " Algebraic foundations of many-valued reasoning", Kluwer Academic, Dordrecht, (2000).
[3] Di Nola, A., Liguori, F. Sessa, S. " Using maximal ideals in the classification of MV-algebras", Prortugaliae Mathematica, vol. 50. Fasc. 1 (1993), 87-102.
[4] Hoo, C. S." MV-algebras, ideals and semisimplicity", Math. Japon., 34, no. 4 (1989), 563-583.
[5] Filipoiu, A., Georgescu, G. Lettieri, A." Maximal MV-algebras", Mathware, soft computing", 4 (1997), 53-62.
[6] Forouzesh, F., Sajadian, F. Bedrood, M. " Inverse topology in MV-algebras", Mathematica Bohemica, No. 3 (2019), 273–285.
[7] Hajek, P. "Metamathematics of fuzzy logic", Kluwer Academic Publishers, Dordrecht 1998.
[8] Iorgulescu, A. " Algebras of logic as BCK-algebras", Academy of economic studies Bucharest, Romania, (2008).
[9] Mundici, D. " Interpretation of AFC^*-algebras in Łukasiewicz sentential calculus", J. Funct. Anal. 65 (1986), 15-63.
[10] Piciu, D. "Algebras of fuzzy logic", Ed. Universitaria Craiova (2007).

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Solving fuzzy bilevel linear programming problem based on interval approximation

1st Niloofar Davoudi, 2nd Farhad Hamidi and 3rd Hassan Mishmast Nehi

1st Faculty of Mathematics. University of Sistan and Baluchestan, Zahedan, Iran, davoudiniloufar@pgs.usb.ac.com
2nd Faculty of Mathematics. University of Sistan and Baluchestan, Zahedan, Iran, f _ hamidi@math.usb.ac.ir
3rd Faculty of Mathematics. University of Sistan and Baluchestan, Zahedan, Iran, hmnehi@hamoon.usb.ac.ir

***Abstract*— In this paper, we propose a method for solving fuzzy bilevel linear programming problem (BLPP) that all parameters are triangular fuzzy numbers (TFNs). First, we transform fuzzy BLPP to interval BLPP. In the next step, we present an algorithm that by using it, the solution of fuzzy BLPP is obtained as TFN.**

***Index Terms*— Bilevel linear programming problem, Fuzzy Numbers, Interval approximation**

I. Introduction

Stackelberg in 1952 introduced BLPP [17]. Then Bracken and McGill [5] formulated it.
In 1965, Zadeh presented theory of fuzzy sets [18].
Lai et al. in [13, 16] solved BLPP with fuzzy method. Zhang et al. [19–22] presented techniques for solving fuzzy BLPP.
Calvete et al. [6] consider BLPP when the coefficients were interval and solving it. Hamidi and Mishmast Nehi [15] generalized the algorithm provided by Calvete et al. [6] for the BLPP. Li and Fang [14] solved interval BLPP by using heuristic algorithm. Davoudi et al. [7] solved fuzzy BLPP by using interval BLPP.
In the next section, we introduce the basic concepts fuzzy and BLPP. In section 3, we consider fuzzy BLPP that all the parameters are TFNs. We convert problem to interval BLPP and solve it. Then we solve an example by using the proposed algorithm.

II. Preliminaries

First, we express a brief summary of the basic concepts required for this paper.

A. Fuzzy number

Definition 1 (Fuzzy set)**.** *[18] Let X be a universe set with element x. A fuzzy set A in X is characterized by a membership function $\mu_A(x)$ which associates with each x in X a real number in the interval $[0,1]$.*

Definition 2. *(Fuzzy number)[2] Let A be a fuzzy set in $\mathbb{R}$. A is called a fuzzy number (FN) if*
(a) $\exists x \in \mathbb{R}$ $s.t.$ $\mu_A(x) = 1$,
(b) $\mu_A(\lambda x_1 + (1-\lambda)x_2) \geq \mu_A(x_1) \wedge \mu_A(x_2)$,
(c) μ_A is upper semicontinuous,
(d) set $\{x \in X : \mu_A(x) > 0\}$ is bounded.

Definition 3. *(Ttriangular fuzzy number)[2] A fuzzy number $A = (a, b, c)$ is called a TFN if its membership function $\mu_A(x)$ is defined as following:*

$$\mu_A(x) = \begin{cases} \frac{x-a}{b-a} & a \leq x \leq b \\ \frac{c-x}{c-b} & b \leq x \leq c \\ 0 & o.w. \end{cases}$$

Definition 4. *(α-cut of a fuzzy number)[2] The α-cut of a fuzzy number A is a set defined as $A^\alpha = [A^L(\alpha), A^U(\alpha)]$ where $A^L(\alpha) = inf\{x \in \mathbb{R} | \mu_A(x) \geq \alpha\}$ and $A^U(\alpha) = sup\{x \in \mathbb{R} | \mu_A(x) \geq \alpha\}$.*

Definition 5 (Adrian and Lucian's nearest interval approximation)**.** *[1] Suppose A be a fuzzy number. Adrian and Lucian's nearest interval approximation for A is defined as $I_A = [m, n]$ where*

$$m = \int_0^1 (\alpha + \frac{1}{2}) A^L(\alpha) d\alpha + \int_0^1 (-\alpha + \frac{1}{2}) A^U(\alpha) d\alpha,$$

$$n = \int_0^1 (-\alpha + \frac{1}{2}) A^L(\alpha) d\alpha + \int_0^1 (\alpha + \frac{1}{2}) A^U(\alpha) d\alpha.$$

B. BLPP

The general form of the BLPP is as (1):

$$\begin{aligned} &\max_{\mathbf{x}_1 \in X_1} Z = \mathbf{c}_1\mathbf{x}_1 + \mathbf{d}_1\mathbf{x}_2 \\ &\text{s.t.} \max_{\mathbf{x}_2 \in X_2} z = \mathbf{c}_2\mathbf{x}_1 + \mathbf{d}_2\mathbf{x}_2 \end{aligned}$$

$$\begin{aligned} \text{s.t.} \quad & A\mathbf{x}_1 + B\mathbf{x}_2 \leq \mathbf{b}, \\ & \mathbf{x}_1, \mathbf{x}_2 \geq \mathbf{0}, \end{aligned} \tag{1}$$

where $\mathbf{x}_1 \in X_1 \subset \mathbb{R}^n$, $\mathbf{x}_2 \in X_2 \subset \mathbb{R}^m$, $F_1, F_2 : X_1 \times X_2 \to \mathbb{R}^1$, $\mathbf{c}_1, \mathbf{c}_2 \in \mathbb{R}^n$, $\mathbf{d}_1, \mathbf{d}_2 \in \mathbb{R}^m$, $\mathbf{b} \in \mathbb{R}^p$, $A \in \mathbb{R}^{p\times n}$ and $B \in \mathbb{R}^{p\times m}$. In the BLPP, first the leader tries to optimize $\mathbf{c}_1\mathbf{x}_1 + \mathbf{d}_1\mathbf{x}_2$ under some constraints by selecting $\mathbf{x}_1$. Then, the follower tries to optimize $\mathbf{c}_2\mathbf{x}_1 + \mathbf{d}_2\mathbf{x}_2$ by choosing $\mathbf{x}_2$ for the specific value of $\mathbf{x}_1$ [4].

So far, several methods have been presented to solve the BLPP, such as Kth-best algorithm [4], sequential LCP method [11], penalty function method [10], branch and bound algorithm [3, 9] and heuristic algorithms [6] ,[12].

III. Fuzzy BLPP

Consider fuzzy BLPP that all the parameters (the objective function, technological coefficients and the resource values) are TFNs:

$$\begin{aligned} \max_{\mathbf{x}_1 \in X_1} \ & Z = \mathbf{c}_1^T\mathbf{x}_1 + \mathbf{d}_1^T\mathbf{x}_2 \\ \text{s.t.} \ & \max_{\mathbf{x}_2 \in X_2} z = \mathbf{c}_2^T\mathbf{x}_1 + \mathbf{d}_2^T\mathbf{x}_2 \\ & \text{s.t.} \quad \mathbf{A}\mathbf{x}_1 + \mathbf{B}\mathbf{x}_2 \geq \mathbf{b} \end{aligned} \tag{2}$$

Where $\mathbf{c}_1 = (\mathbf{c}_1', \bar{\mathbf{c}}_1, \mathbf{c}_1'')$, $\mathbf{d}_1 = (\mathbf{d}_1', \bar{\mathbf{d}}_1, \mathbf{d}_1'')$, $\mathbf{c}_2 = (\mathbf{c}_2', \bar{\mathbf{c}}_2, \mathbf{c}_2'')$, $\mathbf{d}_2 = (\mathbf{d}_2', \bar{\mathbf{d}}_2, \mathbf{d}_2'')$, $A = (A', \bar{A}, A'')$, $B = (B', \bar{B}, B'')$ and $\mathbf{b} = (\mathbf{b}', \bar{\mathbf{b}}, \mathbf{b}'')$ are TFNs.

For solving (2), first we tansform (2) to IBLP and then we solve this IBLP. To convert (2) to IBLP, we use the following theorem:

Theorem 1. *Let A be a TFN. Adrian and Lucian's nearest interval approximation for $A = (a, b, c)$ is given by $I_A = [m, n]$ where*

$$m = \frac{5}{12}a + \frac{7}{12}b + \frac{1}{12}(c - b),$$
$$n = \frac{7}{12}b + \frac{5}{12}c + \frac{1}{12}(a - b).$$

Now, by using Adrian and Lucian's approximation for (2) we have:

$$\begin{aligned} \max_{\mathbf{x}_1 \in X_1} \ & [\tfrac{5}{12}\mathbf{c}_1' + \tfrac{7}{12}\bar{\mathbf{c}}_1 + \tfrac{1}{12}(\mathbf{c}_1'' - \bar{\mathbf{c}}_1), \tfrac{7}{12}\bar{\mathbf{c}}_1 + \tfrac{5}{12}\mathbf{c}_1'' + \tfrac{1}{12}(\mathbf{c}_1' - \bar{\mathbf{c}}_1)]\mathbf{x}_1 \\ & + [\tfrac{5}{12}\mathbf{d}_1' + \tfrac{7}{12}\bar{\mathbf{d}}_1 + \tfrac{1}{12}(\mathbf{d}_1'' - \bar{\mathbf{d}}_1), \tfrac{7}{12}\bar{\mathbf{d}}_1 + \tfrac{5}{12}\mathbf{d}_1'' + \tfrac{1}{12}(\mathbf{d}_1' - \bar{\mathbf{d}}_1)]\mathbf{x}_2 \\ \text{s.t.} \max_{\mathbf{x}_2 \in X_2} \ & [\tfrac{5}{12}\mathbf{c}_2' + \tfrac{7}{12}\bar{\mathbf{c}}_2 + \tfrac{1}{12}(\mathbf{c}_2'' - \bar{\mathbf{c}}_2), \tfrac{7}{12}\bar{\mathbf{c}}_2 + \tfrac{5}{12}\mathbf{c}_2'' + \tfrac{1}{12}(\mathbf{c}_2' - \bar{\mathbf{c}}_2)]\mathbf{x}_1 \\ & + [\tfrac{5}{12}\mathbf{d}_2' + \tfrac{7}{12}\bar{\mathbf{d}}_2 + \tfrac{1}{12}(\mathbf{d}_2'' - \bar{\mathbf{d}}_2), \tfrac{7}{12}\bar{\mathbf{d}}_2 + \tfrac{5}{12}\mathbf{d}_2'' + \tfrac{1}{12}(\mathbf{d}_2' - \bar{\mathbf{d}}_2)]\mathbf{x}_2 \\ \text{s.t.} \ & [\tfrac{5}{12}A' + \tfrac{7}{12}\bar{A} + \tfrac{1}{12}(A'' - \bar{A}), \tfrac{7}{12}\bar{A} + \tfrac{5}{12}A'' + \tfrac{1}{12}(A' - \bar{A})]\mathbf{x}_1 \\ & + [\tfrac{5}{12}\bar{B} + \tfrac{7}{12}\bar{B} + \tfrac{1}{12}(B'' - \bar{B}), \tfrac{7}{12}\bar{B} + \tfrac{5}{12}B'' + \tfrac{1}{12}(\bar{B} - \bar{B})]\mathbf{x}_2 \\ & \leq [\tfrac{5}{12}\bar{\mathbf{b}} + \tfrac{7}{12}\bar{\mathbf{b}} + \tfrac{1}{12}(\mathbf{b}'' - \bar{\mathbf{b}}), \tfrac{7}{12}\bar{\mathbf{b}} + \tfrac{5}{12}\mathbf{b}'' + \tfrac{1}{12}(\bar{\mathbf{b}} - \bar{\mathbf{b}})] \\ & \mathbf{x}_1, \mathbf{x}_2 \geq \mathbf{0}, \end{aligned} \tag{3}$$

To solve (3), we use the following theorem:

Theorem 2. *[8] Consider IBLP as* (4)*.*

$$\begin{aligned} \min_{\boldsymbol{x}_1 \in X_1} \ & Z = [\check{\boldsymbol{c}}_1, \hat{\boldsymbol{c}}_1]\boldsymbol{x}_1 + [\check{\boldsymbol{d}}_1, \hat{\boldsymbol{d}}_1]\boldsymbol{x}_2 \\ \textit{s.t.} \min_{\boldsymbol{x}_2 \in X_2} \ & z = [\check{\boldsymbol{c}}_2, \hat{\boldsymbol{c}}_2]\boldsymbol{x}_1 + [\check{\boldsymbol{d}}_2, \hat{\boldsymbol{d}}_2]\boldsymbol{x}_2 \\ & \textit{s.t.} \quad [\check{A}, \hat{A}]\boldsymbol{x}_1 + [\check{B}, \hat{B}]\boldsymbol{x}_2 \leq [\check{\boldsymbol{b}}, \hat{\boldsymbol{b}}], \\ & \boldsymbol{x}_1, \boldsymbol{x}_2 \geq \boldsymbol{0}, \end{aligned} \tag{4}$$

Where $\check{\boldsymbol{c}}_1$, $\hat{\boldsymbol{c}}_1$, $\check{\boldsymbol{d}}_1$, $\hat{\boldsymbol{d}}_1$, $\check{\boldsymbol{c}}_2$, $\hat{\boldsymbol{c}}_2$, $\check{\boldsymbol{d}}_2$, $\hat{\boldsymbol{d}}_2$, $\check{A}$, $\hat{A}$, $\check{B}$, $\hat{B}$, $\check{\boldsymbol{b}}$ and $\hat{\boldsymbol{b}}$ are real vectors.

The best and the worst optimal values of the leader objective function of the IBLP are the optimal values of the objective functions of the following problems respectively:

$$\begin{aligned} \min_{\boldsymbol{x}_1 \in X_1} \ & Z = \check{\boldsymbol{c}}_1\boldsymbol{x}_1 + \check{\boldsymbol{d}}_1\boldsymbol{x}_2 \\ \textit{s.t.} \min_{\boldsymbol{x}_2 \in X_2} \ & z = \check{\boldsymbol{c}}_2\boldsymbol{x}_1 + \check{\boldsymbol{d}}_2\boldsymbol{x}_2 \\ & \textit{s.t.} \quad \hat{A}\boldsymbol{x}_1 + \hat{B}\boldsymbol{x}_2 \leq \check{\boldsymbol{b}}, \\ & \boldsymbol{x}_1, \boldsymbol{x}_2 \geq \boldsymbol{0}, \end{aligned} \tag{5}$$

$$\begin{aligned} \min_{\boldsymbol{x}_1 \in X_1} \ & Z = \hat{\boldsymbol{c}}_1\boldsymbol{x}_1 + \hat{\boldsymbol{d}}_1\boldsymbol{x}_2 \\ \textit{s.t.} \min_{\boldsymbol{x}_2 \in X_2} \ & z = \hat{\boldsymbol{c}}_2\boldsymbol{x}_1 + \hat{\boldsymbol{d}}_2\boldsymbol{x}_2 \\ & \textit{s.t.} \quad \check{A}\boldsymbol{x}_1 + \check{B}\boldsymbol{x}_2 \leq \hat{\boldsymbol{b}}, \\ & \boldsymbol{x}_1, \boldsymbol{x}_2 \geq \boldsymbol{0}. \end{aligned} \tag{6}$$

By using the Theorem 2, (3) can be divided into the worst and the best subproblems, respectively:

$$\begin{aligned} \max_{\mathbf{x}_1 \in X_1} \ & \big)\tfrac{5}{12}\mathbf{c}_1' + \tfrac{7}{12}\bar{\mathbf{c}}_1 + \tfrac{1}{12}(\mathbf{c}_1'' - \bar{\mathbf{c}}_1)^{\prime}\mathbf{x}_1 + \big)\tfrac{5}{12}\mathbf{d}_1' + \tfrac{7}{12}\bar{\mathbf{d}}_1 + \tfrac{1}{12}(\mathbf{d}_1'' - \bar{\mathbf{d}}_1)^{\prime}\mathbf{x}_2 \\ \text{s.t.} \max_{\mathbf{x}_2 \in X_2} \ & \big)\tfrac{5}{12}\bar{\mathbf{c}}_2 + \tfrac{7}{12}\bar{\mathbf{c}}_2 + \tfrac{1}{12}(\mathbf{c}_2'' - \bar{\mathbf{c}}_2)^{\prime}\mathbf{x}_1 + \big)\tfrac{5}{12}\mathbf{d}_2' + \tfrac{7}{12}\bar{\mathbf{d}}_2 + \tfrac{1}{12}(\mathbf{d}_2'' - \bar{\mathbf{d}}_2)^{\prime}\mathbf{x}_2 \\ & \text{s.t.} \quad \big)\tfrac{7}{12}\bar{A} + \tfrac{5}{12}A'' + \tfrac{1}{12}(A' - \bar{A})^{\prime}\mathbf{x}_1 + \big)\tfrac{7}{12}\bar{B} + \tfrac{5}{12}B'' + \tfrac{1}{12}(B' - \bar{B})^{\prime}\mathbf{x}_2 \\ & \leq \tfrac{5}{12}\bar{\mathbf{b}} + \tfrac{7}{12}\bar{\mathbf{b}} + \tfrac{1}{12}(\mathbf{b}'' - \bar{\mathbf{b}}) \\ & \mathbf{x}_1, \mathbf{x}_2 \geq \mathbf{0} \end{aligned} \tag{7}$$

and

$$\begin{aligned} \max_{\mathbf{x}_1 \in X_1} \ & \big)\tfrac{7}{12}\bar{\mathbf{c}}_1 + \tfrac{5}{12}\mathbf{c}_1'' + \tfrac{1}{12}(\bar{\mathbf{c}}_1 - \bar{\mathbf{c}}_1)^{\prime}\mathbf{x}_1 + \big)\tfrac{7}{12}\bar{\mathbf{d}}_1 + \tfrac{5}{12}\mathbf{d}_1'' + \tfrac{1}{12}(\mathbf{d}_1' - \bar{\mathbf{d}}_1)^{\prime}\mathbf{x}_2 \\ \text{s.t.} \max_{\mathbf{x}_2 \in X_2} \ & \big)\tfrac{7}{12}\bar{\mathbf{c}}_2 + \tfrac{5}{12}\mathbf{c}_2'' + \tfrac{1}{12}(\bar{\mathbf{c}}_2 - \bar{\mathbf{c}}_2)^{\prime}\mathbf{x}_1 + \big)\tfrac{7}{12}\bar{\mathbf{d}}_2 + \tfrac{5}{12}\mathbf{d}_2'' + \tfrac{1}{12}(\mathbf{d}_2' - \bar{\mathbf{d}}_2)^{\prime}\mathbf{x}_2 \\ & \text{s.t.} \quad \big)\tfrac{5}{12}A' + \tfrac{7}{12}\bar{A} + \tfrac{1}{12}(A'' - \bar{A})^{\prime}\mathbf{x}_1 + \big)\tfrac{5}{12}B' + \tfrac{7}{12}\bar{B} + \tfrac{1}{12}(B'' - \bar{B})^{\prime}\mathbf{x}_2 \\ & \leq \big)\tfrac{7}{12}\bar{\mathbf{b}} + \tfrac{5}{12}\mathbf{b}'' + \tfrac{1}{12}(\bar{\mathbf{b}} - \bar{\mathbf{b}})^{\prime} \\ & \mathbf{x}_1, \mathbf{x}_2 \geq \mathbf{0}. \end{aligned} \tag{8}$$

Using the topics discussed above, our proposed algorithm for (2) is expressed as follows:

1. Solve (9) to gain $\bar{\mathbf{x}}_{1opt}$, $\bar{\mathbf{x}}_{2opt}$, $\bar{Z}_{opt}$ and $\bar{z}_{opt}$:

$$\begin{aligned} \max_{\mathbf{x}_1 \in X_1} \ & \bar{\mathbf{c}}_1 x_1 + \bar{\mathbf{d}}_1 x_2 \\ \text{s.t.} \max_{\mathbf{x}_2 \in X_2} \ & \bar{\mathbf{c}}_2 x_1 + \bar{\mathbf{d}}_2 x_2 \\ & \text{s.t.} \bar{A}x_1 + \bar{B}x_2 \leq \bar{\mathbf{b}} \\ & \mathbf{x}_1, \mathbf{x}_2 \geq \mathbf{0}, \end{aligned} \tag{9}$$

(9) is a BLPP that can easily be solved and $\bar{\mathbf{x}}_{1opt}$, $\bar{\mathbf{x}}_{2opt}$, $\bar{Z}_{opt}$ and $\bar{z}_{opt}$ obtain.

2. By solving (7), get $\mathbf{x}'_{1opt}$, $\mathbf{x}'_{2opt}$, Z'_{opt} and z'_{opt}.

3. To obtain $\mathbf{x}''_{1opt}$, $\mathbf{x}''_{2opt}$, Z''_{opt} and z''_{opt}, solve (8).

IV. EXAMPLE

Consider the following problem:

$$\begin{aligned} \max_{x_1 \in X_1} \; & Z = c_1^T x_1 + d_1^T x_2 \\ & \max_{x_2 \in X_2} \; z = c_2^T x_1 + d_2^T x_2 \\ & \text{s.t.} \quad A x_1 + B x_2 \leq b \end{aligned} \tag{10}$$

Where $c_1 = (1, 2, 2.5)$, $d_1 = (0.5, 1.5, 3)$, $c_2 = (3, 4, 6)$, $d_2 = (1.5, 3, 5)$, $A_1 = (1.5, 2, 3)$, $B_1 = (1, 3, 4)$ $b_1 = (6, 10, 13)$ $A_2 = (2, 3, 3.5)$, $B_2 = (0.5, 2, 3)$ and $b_2 = (6, 9, 11)$ are TFNs.
Using the proposed algorithm for problem (10), we will have the following steps:
1.

$$\begin{aligned} \max_{x_1 \in X} \; & 2x_1 + 1.5x_2 \\ & \text{s.t.} \; \max_{x_2 \in Y} \; 4x_1 + 3x_2 \\ & \quad \text{s.t.} \; 2x_1 + 3x_2 \leq 10 \\ & \quad 3x_1 + 2x_2 \leq 9 \\ & \quad x_1, x_2 \geq 0 \end{aligned} \tag{11}$$

That by solvig (11), $\bar{x}_{1opt} = 1.4$, $\bar{x}_{2opt} = 2.4$, $\bar{Z}_{opt} = 6.4$ and $\bar{z}_{opt} = 12.8$ obtain.
By using the Adrian and Lucian's nearest interval approximation, we have:

$$\begin{aligned} \max_{x_1 \in X} \; & [\frac{39}{24}, \frac{51}{24}]x_1 + [\frac{29}{24}, \frac{49}{24}]x_2 \\ & \text{s.t.} \; \max_{x_2 \in Y} \; [\frac{90}{24}, \frac{114}{24}]x_1 + [\frac{61}{24}, \frac{89}{24}]x_2 \\ & \quad \text{s.t.} \; [\frac{45}{24}, \frac{57}{24}]x_1 + [\frac{54}{24}, \frac{78}{24}]x_2 \leq [\frac{206}{24}, \frac{262}{24}] \\ & \quad [\frac{63}{24}, \frac{75}{24}]x_1 + [\frac{35}{24}, \frac{55}{24}]x_2 \leq [\frac{190}{24}, \frac{230}{24}] \\ & \quad x_1, x_2 \geq 0 \end{aligned} \tag{12}$$

Thus, we have:
2.

$$\begin{aligned} \max_{x_1 \in X} \; & \frac{51}{24}x_1 + \frac{49}{24}x_2 \\ & \text{s.t.} \; \max_{x_2 \in Y} \; \frac{114}{24}x_1 + \frac{89}{24}x_2 \\ & \quad \text{s.t.} \; \frac{45}{24}x_1 + \frac{54}{24}x_2 \leq \frac{262}{24} \\ & \quad \frac{63}{24}x_1 + \frac{35}{24}x_2 \leq \frac{230}{24} \\ & \quad x_1, x_2 \geq 0 \end{aligned} \tag{13}$$

By solving (13), we have:
$x'_{1_{opt}} = 1.778$, $x'_{2_{opt}} = 3.369$, $Z'_{opt} = 10.659$ and $z'_{opt} = 20.944$.
3.

$$\begin{aligned} \max_{x_1 \in X} \; & \frac{39}{24}x_1 + \frac{29}{24}x_2 \\ & \text{s.t.} \; \max_{x_2 \in Y} \; \frac{90}{24}x_1 + \frac{61}{24}x_2 \\ & \quad \text{s.t.} \; \frac{57}{24}x_1 + \frac{78}{24}x_2 \leq \frac{206}{24} \\ & \quad \frac{75}{24}x_1 + \frac{55}{24}x_2 \leq \frac{190}{24} \\ & \quad x_1, x_2 \geq 0 \end{aligned} \tag{14}$$

By solving the (14), $x''_{1_{opt}} = 1.285$, $x''_{2\,opt} = 1.701$, $Z''_{opt} = 4.145$ and $z''_{opt} = 9.145$ are achieved.
Thus, $x_{1opt} = (1.285, 1.4, 1.778)$, $x_{2opt} = (1.701, 2.4, 3.369)$, $Z_{opt} = (4.145, 6.4, 10.659)$ and $z_{opt} = (9.145, 12.8, 20.944)$.

V. CONCLUTION

In this paper, a method is proposed for solving special mode of fuzzy BLPP. A simple method that, by calculating and solving three BLPP, obtains the solution as TFNs. We propose that is used this approximation for other type of fuzzy numbers.

REFERENCES

[1] I. B. Adrian, C. Lucian, "Nearest interval, triangular and trapezoidal approximation of a fuzzy," International Journal of Approximate Reasoning, vol. 53, 2012, pp. 805–836.

[2] C. R. Bektor, S. chandra,"Fuzzy Mathematical Programming and Fuzzy Matrix Games," Springer, 2005.

[3] J. Bard, J. Moore, "A branch and bound algorithm for the bilevel programming problem," SIAM J. Sci. Stat. Comput., vol. 11, 1990, pp. 281–292.

[4] W. F. Bialas, , M. H. Karwan, "On two-level optimization," IEEE Transactions on Automatic Control, vol. 27, 1982, pp. 211–214.

[5] J. Bracken, J. McGill, "Mathematical programs with optimization problems in the constraints," Operations Research, vol. 21, 1973, pp. 37–44.

[6] H. I. Calvete, C. Gal, "Linear bilevel programming with interval coefficients," Journal of Computational and Applied Mathematics, vol. 236, 2012, pp. 3751–3762.

[7] N. Davoudi, F. Hamidi and H. Mishmast Nehi, "Triangular fuzzy bilevel linear programming problem," In: 8th Iranian Joint Congress on Fuzzy and Intelligent Systems, CFIS, 2020.

[8] F. Hamidi and H. Mishmast Nehi, "Bilevel Linear Programming with Fuzzy Parameters," Iranian Journal of Fuzzy Systems, vol. 10, 2013, pp. 83–99.

[9] P. Hansen, B. Jaumard, G. Savard, "New branch and bound rules for linear bilevel programming," SIAM Journal on Scientific and Statistical Computing, vol. 13, 1992, pp. 194–217.

[10] Y. T. Hu. Lv, G. Wang, Z. Wan, "A penalty function method based on Kuhn-Tucker condition for solving linear bilevel programming," Applied mathematics and computation ,2007.

[11] J. Judice, A. Faustino, “A sequential LCP method for bilevel linear programming, Annals of Operations Research,” vol. 34, 1992, pp. 89–106.
[12] R. J. Kuo, Y. S. Han, “A hybrid of genetic algorithm and particle swarm optimization for solving bi-level linear programming problem Acase study of supply chain model,” Applied Mathematical Modelling, 2011.
[13] LY. J. ai, ”Hierarchical optimization: A satisfactory solution, “Fuzzy Sets and Systems, vol. 77, 1996, pp. 321–335.
[14] H. Li,L. Fang, “An efficient genetic algorithm for interval linear bilevel programming problems,” Proceedings-9th International Conference on Computational Intelligence and Security, 2013, pp.41–44.
[15] H. Mishmast Nehi, F. Hamidi, , “Upper and lower bounds for the optimal values of the interval bilevel linear programming problem.,” Appl. Math. Model, vol. 39, 2015, pp. 1650–1664.
[16] H. S. Shih, Y. J. Lai, , E. S. Lee, “Fuzzy approach for multilevel programming problems,” Computers and Operations Research, vol. 23, 1996, pp. 73–91.
[17] H. V. Stackelberg, “Theory of the Market Economy,” Oxford University Press, New York, 1952.
[18] L.A. Zadeh, “Fuzzy sets,” Inf. Control, vol. 8(3), 1965, pp. 338–353.
[19] G. Zhang, J. Lu, J. Dillon, “Decentralized multi-objective bilevel decision making with fuzzy demonds,” Base. Syst, vol. 20, 2007, pp. 495–507.
[20] G. Zhang, J. Lu, “Fuzzy bilevel programming with multiple objectives and cooperative multiple followers,” Glob. Optim, vol. 47, 2010, pp. 403–419.
[21] G. Zhang, J. Lu, “The definition of optimal solution and an extended Kuhn-Tucker approach for fuzzy linear bilevel programming,” IEEE Inteligent Information Bulletin, 2005, pp. 6–17.
[22] G. Zhang, J. Lu, J. Dillon, “Fuzzy linear bilevel optimization: Solution concepts, approaches and applications,” Studies in Fuzziness and Soft Computing, vol. 215, 2007, pp. 351–379.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Some Properties of L-graphs

1st Mohammad Mehdi Zahedi and 2nd Elham Raisi Sarbizhan

1st Department of Mathematics, Kerman Graduate University of Advanced Technology, Kerman, Iran, zahedi_mm@kgut.ac.ir
2nd Department of Mathematics, Kerman Graduate University of Advanced Technology, Kerman, Iran, e.raisi@student.kgut.ac.ir

***Abstract*—This paper introduces a graph built on a residual lattice called the L graph. It determines the notion of the strong L-graph. It introduces new concepts, such as direct sum and direct product of two L-graphs. The L-graphs have many applications. For instance, they help determine the minimum number of hospitals equipped with all the treatment wards needed. Accordingly, some related theorems have been proven and several examples have been provided to illustrate these new concepts.**

***Index Terms*—L-graph, strong L-graph, direct product of two L-graphs**

I. Introduction

With the introduction of the concept of the graph by Leonhard Euler, many researchers employed it to model problems in various fields, such as urban planning and life sciences. To date, many concepts have been developed concerning graphs that contributed to the development of this theory. In addition, these new concepts had many applications. In 2008, a group called the AIM Minimum Rank–Special Graphs Work group introduced the concept of forcing set [1]. They used this notion to bound the minimum rank of a graph. Since then, many attempts have been made to investigate this concept, and many have obtained quite fruitful results [5][2]. The fuzzy concept was introduced in 1965 by Zadeh to model uncertain concepts [14]. Many researchers have considered this concept to solve real uncertain problems. For instance, Rosenfeld [11] and Kaufman [6] introduced this concept into graph theory and called it a fuzzy graph. Since then, various problems have been modeled with these types of graphs. Accordingly, in 2021, Mordeson et al. examined several applications in detail in their book. The residuated lattice was introduced by Morgan Ward and Robert P. Dilworth in 1939 [1]. This concept was considered by many researchers and was used in many different branches of science. Recently, Zahedi and Raisi Sarbizhan, etc defined the concept of a graph on a residuated lattice [9][10]. Here are some uses for this type of graph (L-graph), such as determining the minimum number of hospitals in an area to have all the necessary specializations and can be combined. This article presented new concepts such as a direct sum of L-graphs and a direct product of L-graphs, and examples and theorems for these concepts were provided.

II. Preliminaries

This section states the basic definitions of the graph [13][4] and the residuated lattice [12][3] and fuzzy graph [8].

Definition 1. *[1]*

- *Color-change rule:*
  If G is a graph with each vertex colored either white or black, u is a black vertex of G, and exactly one neighbor v of u is white, then change the color of v to black.
- *Given a coloring of G, the derived coloring is the result of applying the color-change rule until no more changes are possible.*
- *A zero forcing set for a graph G is a subset of vertices Z such that if initially the vertices in Z are colored black and the remaining vertices are colored white, the derived coloring of G is all black.*
- *$Z(G)$ is the minimum of $|Z|$ over all zero forcing sets $Z \subseteq V(G)$.*

Definition 2. *[8] $G = (\alpha, \beta)$ is called a fuzzy graph on $G^* = (V, E)$ if $\alpha : V \to [0, 1]$ and $\beta : E \to [0, 1]$ are functions, with $\beta(st) \leq \alpha(s) \wedge \alpha(t)$ for every $st \in E$, where $\wedge$ denotes minimum.*

Definition 3. *[12] A residuated lattice is an algebra $L = (L, \wedge, \vee, \otimes, \to, 0, 1)$ such that*

1) *$L = (L, \wedge, \vee, 0, 1)$ is a lattice (the corresponding order will be denoted by $\leq$) with the smallest element 0 and the greatest element 1,*
2) *$L = (L, \otimes, 1)$ is a commutative monoid (i.e. $\otimes$ is commutative, associative, and $x \otimes 1 = x$ holds),*
3) *$x \otimes y \leq z$ if and only if $x \leq y \to z$ holds (adjointness condition).*

Proposition 1. *[3] Let $(L, \wedge, \vee, \otimes, \to, 0, 1)$ be a residuated lattice. Then the following properties hold:*
(R_1) $1 * x = x$, *where* $* \in \{\wedge, \otimes, \to\}$,
(R_2) $x \otimes 0 = 0, 1' = 0, 0' = 1$,
(R_3) $x \otimes y \leq x \wedge y \leq x, y$, *and* $y \leq (x \to y)$,
(R_4) $(x \to y) \otimes x \leq y$,

(R_5) $x \leq y$ *implies* $x * z \leq y * z$, *where* $* \in \wedge, \vee, \otimes$,
(R_6) $z \otimes (x \wedge y) \leq (z \otimes x) \wedge (z \otimes y)$,
(R_7) $x \otimes (y \vee z) = (x \otimes y) \vee (x \otimes z)$.

Definition 4. *[9] Let* $G_1 = (\alpha_1, \beta_1)$ *and* $G_2 = (\alpha_2, \beta_2)$ *be two L-graphs on* $G_1^* = (V_1, E_1)$ *and* $G_2^* = (V_2, E_2)$, *respectively, and* $c \in L \backslash \{1\}$. *Then* G_1 *and* G_2 *are isomorphic with threshold c, denoted by* $G_1 \cong G_2$ *if there exists a bijection* h *from* V_1 *into* V_2 *such that the following conditions hold true for all* $u, v \in V_1$*:*

(i) $uv \in E_1$ *if and only if* $h(u)h(v) \in E_2$,
(ii) $\alpha_1(u) > c$ *if and only if* $\alpha_2(h(u)) > c$,
(iii) $\beta_1(uv) > c$ *if and only if* $\beta_2(h(u)h(v)) > c$.

h *is an isomorphism* $(\cong)$ *if and only if* h *is an isomorphism with threshold* c *for every* $c \in L \backslash \{1\}$.

III. L-Graphs

This section notifies the notion of an L-graph built on a residuated lattice. In this paper, it was assumed that L is a residuated lattice, and G^* is a simple graph (V, E).

The fuzzy set was considered a residuated lattice. Thus, some definitions in our work were provided by the residuated lattice. The L-graph built on the residuated lattice was defined using the fuzzy graph definition [7]. Also, the direct sum and direct product were expanded, and several examples were provided to clarify the concepts. Besides, a strong L-graph was determined, proving that if two L-graphs are strong, their direct sum and their direct product would also be strong. An example showed that two L-graphs might not be strong, but their direct product is strong. Then, it was also shown that the two isomorphic L-graphs were strong if and only if their direct product was strong. It is shown that if at least one of the two L-graphs is disconnected L-graph, then their direct product is disconnected L-graph.

Definition 5. $G = (\alpha, \beta)$ *is called an L-graph on* G^* *if* $\alpha : V \to L$ *and* $\beta : E \to L$ *are functions, with* $\beta(st) \leq \alpha(s) \otimes \alpha(t)$ *for every* $st \in E$. *Besides, if* G^* *is a path (cycle, bipartite, complete, complete bipartite) graph, then* G *is called a path (cycle, bipartite, complete, complete bipartite) L-graph on* G^*.

Example 1. *Suppose* $L = ([0,1], \wedge, \vee, \otimes, \rightarrow, 0, 1)$, *where* $a \otimes b = \begin{cases} (a+b-1) & if\ a+b \geq 1, \\ 0 & if\ a+b < 1, \end{cases}$ *and* $a \rightarrow b = \begin{cases} 1 & if\ b-a \geq 0, \\ (1-a+b) & if\ b-a < 0. \end{cases}$ *So,* $G = (\alpha, \beta)$ *is a cycle L-graph on* G^*, *as in Fig. 1, where* $V = \{u_1, u_2, u_3\}$, $E = \{u_1u_2, u_2u_3, u_1u_3\}$, $\beta(u_iu_j) = \alpha(u_i) \otimes \alpha(u_j)$, $\alpha(u_1) = 0.9$, $\alpha(u_2) = 0.7$, $\alpha(u_3) = 0.2$, $\beta(u_1q_2) = 0.6$, $\beta(u_2q_3) = 0$, *and* $\beta(u_1u_3) = 0.1$. *Thus,* $Z(G) = \{u_1, u_2\}$ *is one of the zero forcing sets of G, as in Fig. 1.*

Definition 6. *Let* $G = (\alpha, \beta)$ *be an L-graph on* G^* *such that* $\beta(st) = \alpha(s) \otimes \alpha(t)$, *for every* $st \in E$. *Then, G is a strong L-graph.*

Definition 7. *Consider* L *and two L-graphs* $G_1 = (\alpha_1, \beta_1)$ *and* $G_2 = (\alpha_2, \beta_2)$ *on* $G_1^* = (V_1, E_1)$ *and* $G_2^* = (V_2, E_2)$, *respectively.*

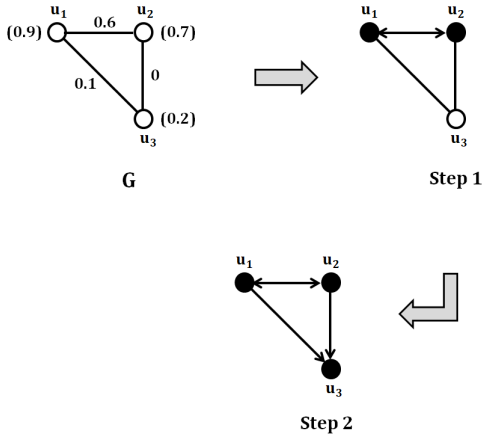


Fig. 1. The L-graph G, and $Z(G) = \{u_1, u_2\}$.

(i) *Then,* $G = G_1 + G_2 = (\alpha, \beta)$ *on* $G^* = (V, E)$ *is a direct sum of two L-graphs, where* $V_1 \cap V_2 = \emptyset$, $V = V_1 \cup V_2$, $E = E_1 \cup E_2$, $\beta(uu') = \begin{cases} \beta_1(uu') & if\ uu' \in E_1, \\ \beta_2(uu') & if\ uu' \in E_2, \end{cases}$ *and* $\alpha(u) = \begin{cases} \alpha_1(u) & if\ u \in V_1, \\ \alpha_2(u) & if\ u \in V_2. \end{cases}$

(ii) *Then,* $G' = G_1 \times G_2 = (\alpha', \beta')$ *on* $G'^* = (V', E')$ *is a product of two L-graphs, where* $V' = V_1 \times V_2$, $(u, u')(v, v') \in E'$ *if and only if* $uv \in E_1$ *and* $u'v' \in E_2$, $\beta'((u, u')(v, v')) = \beta_1(uv) \otimes \beta_2(u'v')$, *and* $\alpha'((u, u')) = \alpha_1(u) \otimes \alpha_2(u')$, *for every* $u \in V_1$, *and* $u' \in V_2$.

Theorem 1. *Let* $G_1 = (\alpha_1, \beta_1)$ *and* $G_2 = (\alpha_2, \beta_2)$ *be two L-graphs on* $G_1^* = (V_1, E_1)$, *and* $G_2^* = (V_2, E_2)$, *respectively. Then,*

(i) *If G is their direct sum, then it is the L-graph,* $|Z(G)| = |Z(G_1)| + |Z(G_2)|$, *and* $Z(G) = Z(G_1) \cup Z(G_2)$ *is one of the zero forcing sets of G.*
(ii) *If* G' *is their direct product, then it is the L-graph.*
(iii) G_1 *and* G_2 *are strong L-graphs G if and only if their direct sum is a strong L-graph.*
(v) *If* G_1 *and* G_2 *are strong L-graphs, then their direct product* G' *is a strong L-graph.*

Proof: (i) Using the definition of the direct sum of two L-graphs, the proof is straight forward.

(ii) To prove that $G' = (\alpha', \beta')$ *on* $G'^* = (V', E')$ *is an L-graph is sufficient, show that* $\beta'((u, u')(v, v')) \leq \alpha'((u, u')) \otimes \alpha'((v, v'))$, *for every* $(u, u')(v, v') \in E'$.

So,

$$\begin{aligned} \beta'((u,u')(v,v')) &= \beta_1(uv) \otimes \beta_2(u'v'), \text{by definition of } \beta' \\ &\leq (\alpha_1(u) \otimes \alpha_2(u')) \otimes (\alpha_1(v) \otimes \alpha_2(v')) \\ &\leq \alpha'((u,u')) \otimes \alpha'((v,v')), \\ &\quad \text{by definition of } \alpha', \end{aligned}$$

for every $(u, u')(v, v') \in E'$.

(iii) By definition of direct sum, it is proved.

(v) Assume that $G' = (\alpha', \beta')$ on $G'^* = (V', E')$ is their direct product. So,

$$\begin{aligned}\beta'((u,u')(v,v')) &= \beta_1(uv) \otimes \beta_2(u'v'),\\ &= (\alpha_1(u) \otimes \alpha_2(u')) \otimes (\alpha_1(v) \otimes \alpha_2(v')),\\ &\quad \textit{by definition of strong L-graph}\\ &= \alpha'((u,u')) \otimes \alpha'((v,v')),\\ &\quad \textit{by definition of } \alpha',\end{aligned}$$

for every $(u,u')(v,v') \in E'$. Thus, it is the strong L-graph. ■

Example 2. *Consider $L = ([0,1], \wedge, \vee, \otimes, \rightarrow, 0, 1)$, where $a \otimes b = a \wedge b$, and $a \rightarrow b = \begin{cases} 1 & if\ a \leq b,\\ \frac{b}{a} & if\ otherwise,\end{cases}$ and two L-graphs $G_1 = (\alpha_1, \beta_1)$ on $G_1^* = (V_1, E_1)$ and $G_2 = (\alpha_2, \beta_2)$ on $G_2^* = (V_2, E_2)$, as in Fig. 2, where $V_1 = \{u_1, u_2, u_3\}$, $E_1 = \{u_1u_2, u_2u_3\}$ $\alpha_1(u_1) = 0.7$, $\alpha_1(u_2) = 0.8$, $\alpha_1(u_3) = 0.5$, $\beta_1(u_1u_2) = 0.7$, $\beta_1(u_2u_3) = 0.5$, $V_2 = \{u_1', u_2'\}$, $E_2 = \{u_1'u_2'\}$, $\alpha_2(u_1') = 0.9$, $\alpha_2(u_2') = 1$, $\beta_2(u_1'u_2') = 0.8$, $Z(G_1) = \{u_1\}$, and $Z(G_2) = \{u_1'\}$. So, $G = (\alpha, \beta)$ on $G^* = (V, E)$ is their direct sum, as in Fig. 2, where $V = V_1 \cup V_2$, $E = E_1 \cup E_2$, $\alpha(u_1) = 0.7$, $\alpha(u_2) = 0.8$, $\alpha(u_3) = 0.5$, $\alpha(u_1') = 0.9$, $\alpha(u_2') = 1$, $\beta(u_1u_2) = 0.7$, $\beta(u_2u_3) = 0.5$, $\beta(u_1'u_2') = 0.8$, and $Z(G) = \{u_1, u_1'\}$, as in Fig. 3. Also, $G' = (\alpha', \beta')$ on $G'^* = (V', E')$ is their direct product, as in Fig. 4, where $V' = \{(u_1,u_1'), (u_1,u_2'), (u_2,u_1'), (u_2,u_2'), (u_3,u_1'), (u_3,u_2')\}$, $E' = \{(u_1,u_1')(u_2,u_2'), (u_1,u_2')(u_2,u_1'), (u_2,u_1')(u_3,u_2'), (u_3,u_1')(u_2,u_2')\}$, $\alpha'((u_1,u_1')) = 0.7$, $\alpha'((u_1,u_2')) = 0.7$, $\alpha'((u_2,u_1')) = 0.8$, $\alpha'((u_2,u_2')) = 0.8$, $\alpha'((u_3,u_1')) = 0.5$, $\alpha'((u_3,u_2')) = 0.5$, $\beta'((u_1,u_1')(u_2,u_2')) = 0.7$, $\beta'((u_1,u_2')(u_2,u_1')) = 0.7$, $\beta'((u_2,u_1')(u_3,u_2')) = 0.5$, and $\beta'((u_2,u_2')(u_3,u_1')) = 0.5$. Hence, $Z(G) = \{(u_1,u_1'), (u_1,u_2')\}$, as in Fig. 4.*

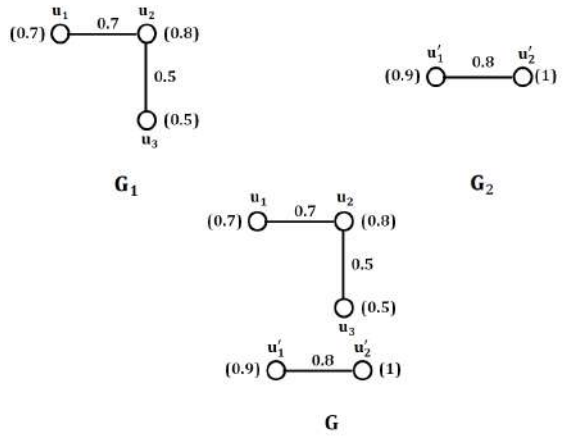


Fig. 2. The L-graph G_1, the L-graph G_2, and the direct sum L-graph G.

Remark 1. *The example above expresses the possibility of having two non-strong L-graphs, including G_1 and G_2, while their direct product is a strong L-graph and G is a disconnected direct product of two L-graphs.*

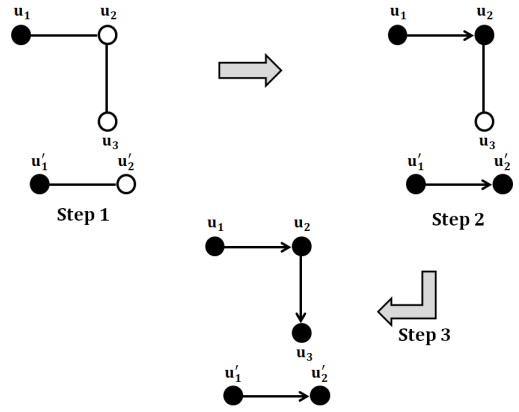


Fig. 3. $Z(G) = \{u_1, u_1'\}$.

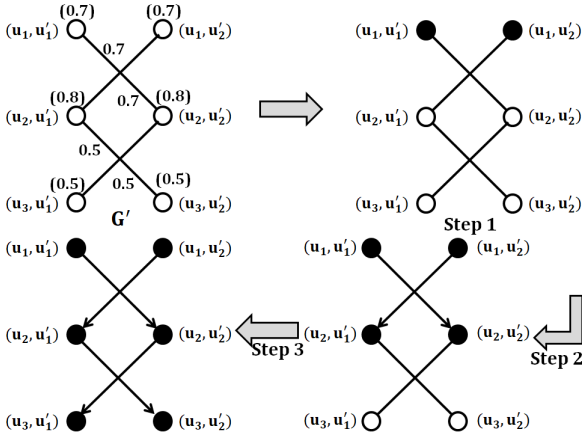


Fig. 4. The direct product L-graph G', and $Z(G') = \{(u_1,u_1'), (u_1,u_2')\}$

Theorem 2. *Suppose two isomorphic L-graphs $G_1 = (\alpha_1, \beta_1)$ on $G_1^* = (V_1, E_1)$, and $G_2 = (\alpha_2, \beta_2)$ on $G_2^* = (V_2, E_2)$. Then they are strong L-graphs if and only if their direct product is a strong L-graph.*

Proof: If G_1 and G_2 are strong L-graphs, by Theorem 1(v), then their direct product is strong L-graph.

If $G' = (\alpha', \beta')$ on $G'^* = (V', E')$ is a strong L-graph, and $h : V_1 \longrightarrow V_2$ is isomorphism, then

$$\begin{aligned}\beta'((u,h(u))(v,h(v))) &= \beta_1(uv) \otimes \beta_2(h(u)h(v)),\\ &= (\alpha_1(u) \otimes \alpha_1(v))\\ &\otimes\ (\alpha_2(h(u)) \otimes \alpha_2(h(v))),\\ &\quad \textit{by definition of strong L-graph.}\end{aligned}$$

Proof by contradiction, consider

$$\beta_1(uv) \neq \alpha_1(u) \otimes \alpha_1(v) \Leftrightarrow \beta_1(uv) < \alpha_1(u) \otimes \alpha_1(v).$$

Since G_1 and G_2 isomorphic L-graphs, $\beta_2(h(u)h(v)) < \alpha_2(h(u)) \otimes \alpha_2(h(v))$. Hence,

$$\begin{aligned}\beta_1(uv) \otimes \beta_2(h(u)h(v)) &< \beta_1(uv) \otimes (\alpha_2(h(u)) \otimes \alpha_2(h(v)))\\ &< (\alpha_1(u) \otimes \alpha_1(v))\\ &\otimes\ (\alpha_2(h(u)) \otimes \alpha_2(h(v))).\end{aligned}$$

So, it contradicts the strong L-graph G'. Then, G_1 and G_2 are strong L-graphs. ■

Theorem 3. *Let G_1 and G_2 be two L-graphs, and at least one of them be disconnected L-graph. Then their direct product is a disconnected L-graph.*

Proof: Consider two L-graphs $G_1 = (\alpha_1, \beta_1)$ and $G_2 = (\alpha_2, \beta_2)$ on $G_1^* = (V_1, E_1)$ and $G_2^* = (V_2, E_2)$, respectively. Also, assume their direct product $G' = (\alpha', \beta')$ on $G'^* = (V', E')$. As $(u, u')(v, v') \in E'$ if and only if $uv \in E_1$ and $u'v' \in E_2$ and by using of the definition of direct product L-graph, it is proved. ■

Example 3. *Assume L, the L-graph G in Example 1, and the L-graph $G_2 = (\alpha_2, \beta_2)$ on $G_2^* = (V_2, E_2)$, as in Fig. 5, where $V_2 = \{u'_1, u'_2, u'_3\}$, $E_2 = \{u'_2u'_2\}$, $\alpha_1(u'_1) = 0.5$, $\alpha_1(u'_2) = 0.8$, $\alpha_1(u'_3) = 0.9$, and $\beta_1(u'_2u'_3) = 0.6$, Hence, $G' = (\alpha', \beta')$ on $G'^* = (V', E')$ is a disconnected direct product of two L-graphs, as in Fig. 5, where $V' = \{(u_1, u'_1), (u_1, u'_2), (u_1, u'_3), (u_2, u'_1), (u_2, u'_2), (u_2, u'_3), (u_3, u'_1), (u_3, u'_2), (u_3, u'_3)\}$, $E' = \{(u_1, u'_2)(u_2, u'_3), (u_1, u'_2)(u_3, u'_3), (u_2, u'_2)(u_1, u'_3), (u_2, u'_2)(u_3, u'_3), (u_3, u'_2)(u_1, u'_3), (u_3, u'_2)(u_2, u'_3)\}$, $\alpha'((u_1, u'_1)) = 0.4$, $\alpha'((u_1, u'_2)) = 0.7$, $\alpha'((u_1, u'_3)) = 0.8$, $\alpha'((u_2, u'_1)) = 0.2$, $\alpha'((u_2, u'_2)) = 0.5$, $\alpha'((u_2, u'_3)) = 0.6$, $\alpha'((u_3, u'_1)) = 0$, $\alpha'((u_3, u'_2)) = 0$, $\alpha'((u_3, u'_3)) = 0.1$, $\beta'((u_1, u'_2)(u_2, u'_3)) = 0.2$, $\beta'((u_1, u'_2)(u_3, u'_3)) = 0$, $\beta'((u_2, u'_2)(u_1, u'_3)) = 0.2$, $\beta'((u_2, u'_2)(u_3, u'_3)) = 0$, $\beta'((u_3, u'_2)(u_1, u'_3)) = 0$, and $\beta'((u_3, u'_2)(u_2, u'_3)) = 0$.*

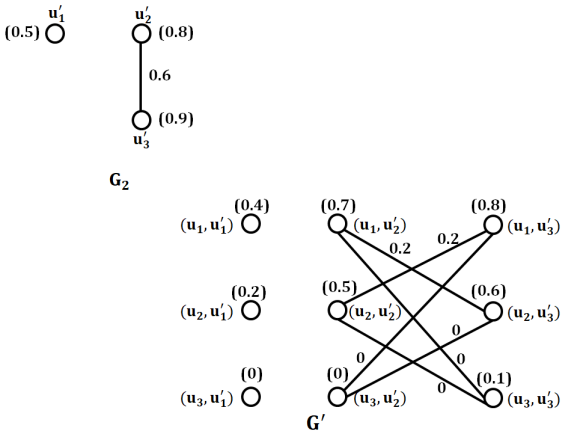


Fig. 5. The L-graph G_2, and the L-graph G'.

Remark 2. *The following example expresses that it is possible to have two connected L-graphs, G_1 and G_2, while their direct product is a disconnected L-graph.*

Example 4. *Consider L and two L-graphs G_1 and G_2 and their direct product L-graph G' in Example 2. Thus, their direct product is a disconnected L-graph, while G_1 and G_2 are connected L-graphs.*

IV. Some applications of L-graphs

In [8], many applications of fuzzy graphs are mentioned by Mordeson and his co-authors, but these graphs are not used in some cases, while the novel graphs, L-graph, can be useful in other applications in addition to those applications. For instance, it can apply to categorizing some hospitals to connect them and determine the minimum number of hospitals that have m treatment wards.

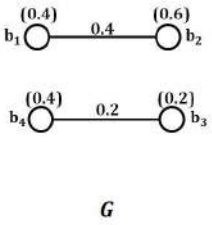


Fig. 6. The fuzzy graph G.

Example 5. *Assume that a given region has four hospitals named b_1, b_2, b_3, and b_4 and five treatment wards, namely a_1, a_2, a_3, a_4, and a_5 such that Hospital b_1 has two treatment wards a_1, and a_2, Hospital b_2 has three treatment wards, including a_1, a_2, and a_3, Hospital b_3 has a treatment ward, namely a_4, and Hospital b_4 has two treatment wards, including a_4, and a_5. Then, $G = (\alpha, \beta)$ is a fuzzy graph on $G^* = (V, E)$, as in Fig. 6, where $V = \{b_1, b_2, b_3, b_4\}$, $b_ib_j \in E$ if and only if these two books have at least one similar topic,*

$$\alpha(b_i) = \frac{\textit{the number of treatment wards of } b_i}{\textit{the number of all treatment wards}},$$

for every $i = 1, 2, 3, 4$, and $\beta(b_ib_j) = \alpha(b_i) \wedge \alpha(b_j)$. Hence, $\alpha(b_1) = 0.4$, $\alpha(b_2) = 0.6$, $\alpha(b_3) = 0.2$, $\alpha(b_4) = 0.4$, $\beta(b_1b_2) = 0.4$, and $\beta(b_3b_4) = 0.2$.

Remark 3. *The example above shows that modeling a library by fuzzy graph has problems such as not seeing the book contents after modeling and not being able to determine similar topics of books after modeling.*

Definition 8. *Let $(L, \wedge, \vee, \otimes, \rightarrow, 0, 1)$, and $G = (V, E)$ be the L-graph. Then $K(G) = \{v_i | v_i \in V, \vee v_i = 1\}$ is the cover set of this L-graph when it has a minimum number of vertices and this L-graph is called the close L-graph.*

Aplication 1. *The L-graphs have some applications. For instance;*

(a:) *Suppose a region has n hospitals that each of them has some treatment wards. Also, we need some hospitals that have m treatment wards. Therefore, the L-graph is used to categorize these hospitals to connect them and determine the minimum hospitals with m treatment wards.*
Assume $L = (P(X), \cap, \cup, \otimes, \rightarrow, \emptyset, X)$, where $X = \{a_1, a_2, \ldots, a_m\}$ so that a_i is labeled per i for m treatment wards of these hospitals and $A \otimes B = A \cap B$, and $A \rightarrow B = \begin{cases} X & if\ A \subseteq B, \\ B & if\ B \subset A. \end{cases}$ Then $G = (\alpha, \beta)$ is an L-graph on $G^* = (V, E)$, where $V = \{b_1, b_2, \ldots, b_n\}$ such that b_is are

labels of these hospitals for each $i = 1, 2, \ldots, n$. Also, $b_ib_j \in E$ if and only if these two hospitals have at least one similar treatment ward, $\alpha(b_i) = \{a_k | a_k$ is one of the treatment wards of $b_i\}$, and $\beta(b_ib_j) = \alpha(b_i) \otimes \alpha(b_j)$, for every $b_ib_j \in E$. In this application, the cover set of the L-graph determines the minimum of hospitals that cover all treatment wards. Besides, in this case, to determine the cover set, one only needs to start selecting the vertices that have the maximum degree. This process is continued until their union covers all the treatment wards.

b: *Suppose that a construction company decides to contract with several sales materials and wants to contract with the least of them so that these material sellers can supply the company's m construction materials. Consider $L = (P(X), \cap, \cup, \otimes, \rightarrow, \emptyset, X)$, where $X = \{a_1, a_2, \ldots, a_m\}$ so that a_i is labeled per i for m construction materials of these sales materials and $A \otimes B = A \cap B$, and $A \rightarrow B = \begin{cases} X & if\ A \subseteq B, \\ B & if\ B \subset A. \end{cases}$ Then $G = (\alpha, \beta)$ is an L-graph on $G^* = (V, E)$, where $V = \{b_1, b_2, \ldots, b_n\}$ such that b_is are labels of these sales materials for each $i = 1, 2, \ldots, n$. Also, $b_ib_j \in E$ if and only if these two sales materials have at least one similar construction material, $\alpha(b_i) = \{a_k | a_k$ is one of the construction materials of $b_i\}$, and $\beta(b_ib_j) = \alpha(b_i) \otimes \alpha(b_j)$, for every $a_ia_j \in E$. In this application, the cover set of the L-graph determines the minimum of sales materials that cover all construction materials. Besides, in this case, to determine the cover set, one only needs to start selecting the vertices that have the maximum degree. This process continues until their union covers all the construction materials.*

Example 6. *Example 6: Assume the region has four hospitals: a_1, a_2, a_3, and a_4. Hospital a_1 has two treatment wards, including neurology and cardiology. Hospital a_2 has three treatment wards, including pediatric, neurology, and general. Hospital a_3 has two treatment wards, including neurology and general, and Hospital a_4 has pediatric and cardiovascular care wards. So, they have four treatment wards: pediatric, neurology, general, and cardiology. Hence, these treatment wards are labeled b_1, b_2, b_3, and b_4, respectively. Then, $G = (\alpha, \beta)$ is an L-graph on $G^* = (V, E)$, as in Fig. 7, where $V = \{a_1, a_2, a_3, a_4\}$, $E = \{a_1a_2, a_1a_3, a_1a_4, a_2a_3, a_2a_4\}$, $\alpha(a_1) = \{b_2, b_4\}$, $\alpha(a_2) = \{b_1, b_2, b_3\}$, $\alpha(a_3) = \{b_2, b_3\}$, $\alpha(a_4) = \{b_1, b_4\}$, $\beta(a_1a_2) = \{b_2\}$, $\beta(a_1a_3) = \{b_2\}$, $\beta(a_1a_4) = \{b_4\}$, $\beta(a_2a_3) = \{b_2, b_3\}$, and $\beta(a_2a_4) = \{b_1\}$. So, the cover sets of G are a_1, and a_2 or a_2, and a_4.*

V. Conclusion

A notion of L-graph is notified in this paper based on a residuated lattice. Accordingly, The notions of strong L-graph, direct sum and direct product are also represented and some theorems and examples are provided. Besides, some applications are shown in L-graph and two of them are described in detail in this paper. The L-graph is utilized to model the thematic relevance of hospitals in a region or sort some treatment wards. Today, one of the most important issues facing modern cities is urban and traffic plans. Likewise, one of the concerns of scientists today is to prevent the spread of the disease across the country. Another of their purpose is the treatment of diseases, which can be solved by using conceptual definitions in this article. Our next sight is to do more research on structures and their complexity. In addition, we intend to take steps to improve the performance of the education system.

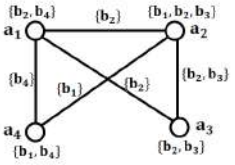


Fig. 7. The L-graph G.

References

[1] AIM Minimum Rank–Special Graphs Work Group, "Zero forcing sets and the minimum rank of graphs,"Linear Algebra and its Applications, vol. 428, no. 7, p.p 1628-1648, Apr. 2008.

[2] F. Barioli, W. Barrett, S. Fallat, H.T. Hall, L. Hogben, B. Shader, P. van den Driessche, and H. van der Holst, "Parameters related to tree-width, zero forcing, and maximum nullity of a graph,"Journal of Graph Theory, vol. 72, no. 2, pp. 146-177, Jun. 2012.

[3] K. Blount, and C. Tsinakis, "The structure of residuated lattices,"International Journal of Algebra and Computation, vol. 13, n. 4, pp. 437– 461, Aug. 2003.

[4] J.A Bondy, and U.S. R. Murty, Graph theory, Landon, UK, Springer London, 2008, pp. 29-86.

[5] L. Deaett, and S. A. Meyer, "The minimum rank problem for circulants,"vol. 491, no. 4, pp. 386-418, Feb. 2016.

[6] A. Kauffman, Introduction a La theorie des sous-ensembles flous: A L'Usage Des Ingenieurs (Fuzzy Sets Theory), Tome III, Paris, French, Masson Et Cie, Editeurs, 1975.

[7] S. Mathew, J.N. Mordeson, and D.S. Malik, Fuzzy graph theory, Springer International Publishing, 2017, pp. 13-116.

[8] J. N. Mordeson, and S. Mathew, Mathematics of Uncertainty for Coping with World Challenges: Climate Change, World Hunger, Modern Slavery, Coronavirus, Human Trafficking, Springer International Publishing, 2021, pp. 133-173.

[9] E. Raisi Sarbizhan, M.M. Zahedi, and M. Sahmsizade, "L-graph automata and some applications,"submitted.

[10] E. Raisi Sarbizhan, and M.M. Zahedi, "Maximal product of two L-graph automata and some applications,"submitted.

[11] A. Rosenfeld, “Fuzzy groups,”Journal of Mathematical Analysis and Applications, vol. 35, no. 3, p.p 512-517, Sep. 1971.
[12] M. Ward, and R.P.Dilworth, “Residuated lattices,”Transactions of The American Society, vol. 45, no. 3, pp. 335–354, Apr. 1939.
[13] D.B. West, Introduction to graph theory, (Vol. 2), Delhi, India, Pearson Education 2001, pp. 19-67.
[14] L.A. Zadeh, “Similarity relations and fuzzy orderings,”Information Sciences, vol. 3, no. 2, p.p 177–200, Apr. 1971.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Some Results in Projective System of *BL*-algebras

Reza Tayebi Khorami[1], Arsham Borumand Saeid[2]

[1]*Department of Mathematics, Ahvaz Branch, Islamic Azad University, Ahvaz, Iran*, r.t.khorami@gmail.com
[2] *Department of Pure Mathematics, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman, Kerman, Iran*
arsham@mail.uk.ac.ir

***Abstract*— In this paper, we introduce the definition of inverse limit of a projective system of *BL*-algebras and study some basic properties. Also, we consider the set of congruences of a *BL*-algebra as the poset, then we construct a quotient inverse system and quotient inverse limit on it.**

***Index Terms* —BL-algebra, inverse limit, projective system.**

I. INTRODUCTION

H'ajek introduced basic logic *BL* as the logic of continuous t-norms and their residua [4]. Basic logic is a fuzzy logic, i.e. it is complete with respect to linearly ordered models. Algebraic semantics of *BL* is the variety of *BL*-algebras. It was proved by Cignoli, Esteva, Godo and Torrens that the variety of *BL*-algebras is generated just by the continuous t-norms on the interval [0, 1] of reals.

The origins of the study of inverse limits date back to the 1920's. Classical theory of inverse systems and inverse limits are important in the extension of homology and cohomology theory. An exhaustive discussion of inverse systems which are in the some classical categories such as **Set, Top, Grp** and **Rng** defined in [1], was presented by the paper [3] which is a milestone in the development of that theory.

As is the case with products, the inverse limit might not exist in any category in general where as inverse systems exist in every category. Note from that [3] inverse limits exist in any category when that category has products of objects and the equalizers [1] of pairs of morphisms, in other words, the inverse limits exist in any category if the category is complete, in the sense of [1].

Additionally, an inverse system has at most one limit. That is, if an inverse limit of any inverse system exists in any category C, this limit is unique up to C-isomorphism. Incidentally, inverse limits always exist in the categories **Set, Top, Grp** and **Rng**. Note also that inverse limits are generally restricted to diagrams over directed sets.

An inverse limit (also called the projective limit) is a construction that allows one to "glue together" several related objects, the precise gluing process being specified by morphisms between the objects. Inverse limits are a special case of the concept of limit in category theory. We will apply the concept of inverse limit in the sense of category theory to some collections of *BL*-algebras. This notion in category theory has been studied in different kinds of categories. In this paper, we introduce the concepts of inverse limit of an inverse system of *BL*-algebras.

II. PRELIMINARIES

In this section, some basic definitions and properties are recalled to be considered later.

Definition 2.1. [4] A *BL*-algebra is an algebra $(A, \vee, \wedge, \odot, \rightarrow, 0, 1)$ of type (2,2,2,2,0,0) such that:

(BL1) $(A, \vee, \wedge, 0, 1)$ is a bounded lattice,
(BL2) $(A, \odot, 1)$ is an abelian monoid,
(BL3) $x \odot z \leq y$ if and only if $z \leq x \rightarrow y$,
(BL4) $x \odot (x \rightarrow y) = x \wedge y$,
(BL5) $(x \rightarrow y) \vee (y \rightarrow x) = 1$ for all $x, y \in A$.

In this paper, we denote *BL*-algebra $(A, \vee, \wedge, \odot, \rightarrow, 0, 1)$ with A.

Definition 2.2. [4] Let A be a *BL*-algebra. A nonempty subset F of A is called a filter of A if F satisfies the following conditions:

(F1) if $x \in F$, $x \le y$ and $y \in A$ then $y \in F$,
(F2) $x \odot y \in F$ for every $x, y \in A$, that is, F is a subsemigroup of A.

A mapping $f : X \longrightarrow Y$ is a *BL*-homomorphism if $f(x \odot y) = f(x) \odot f(y)$ and $f(x \to y) = f(x) \to f(y)$ for all $x, y \in A$.

Definition 2.3. [2] Let I be a set and "$\le$" be a binary operation on I. We call $I = (I, \le)$ a directed partially ordered set or directed poset if it satisfies the following conditions:

(i) $i \le i$, for all $i \in I$,
(ii) $i \le j$ and $j \le i$ imply $i = j$, for $i, j \in I$,
(iii) $i \le j$ and $j \le k$ imply $i \le k$, for $i, j, k \in I$,
(iv) if $i, j \in I$, there exists some $k \in I$ such that $i, j \le k$.

III. PROJECTIVE SYSTEMS

Definition 3.1. An inverse or projective system of *BL*-algebras over a directed poset I, consists of a collection $\{A_i \mid i \in I\}$ of *BL*-algebras indexed by I, and a collection of *BL*-homomorphisms $\{\varphi_{ij} : A_i \longrightarrow A_j\}$, defined whenever $i \ge j$, such that $\varphi_{ii} = id_{A_i}$, $\varphi_{jk}\varphi_{ij} = \varphi_{ik}$ whenever $i, j, k \in I$ and $i \ge j \ge k$.

We denote an inverse system by $\{A_i, \varphi_{ij}, I\}$ or $\{A_i, \varphi_{ij}\}$.

Definition 3.2. Let B be a *BL*-algebra, $\{A_i, \varphi_{ij}, I\}$ an inverse system of *BL*-algebras, and $\psi_i : B \longrightarrow A_i$ be a *BL*-homomorphism for each $i \in I$. These mappings ψ_i are called to be compatible if $\varphi_{ij}\psi_i = \psi_j$ whenever $i \ge j$.

Definition 3.3. Let $\{A_i, \varphi_{ij}, I\}$ be an inverse system of *BL*-algebras. A subalgebra A of $\prod_{i \in I} A_i$ together with compatible homomorphisms $\phi_i : A \longrightarrow A_i$ is an inverse limit or a projective limit of $\{A_i, \varphi_{ij}, I\}$ if the following universal property is satisfied: whenever B is a *BL*-algebra and $\{\psi_i : B \longrightarrow A_i\}$ is a set of compatible homomorphisms, then there is a unique homomorphism $\psi : B \longrightarrow A$ such that $\varphi_i\psi = \psi_i$ for all $i \in I$.

We say that ψ is "induced" by the compatible homomorphisms ψ_i. The maps $\varphi_i : A \longrightarrow A_i$ are called projections. We shall denote the inverse limit of the inverse system $\{A_i, \varphi_{ij}, I\}$ by $\lim_{\leftarrow i \in I} A_i$ or (A, φ_i).

Theorem 3.4. Let $\{A_i, \varphi_{ij}, I\}$ be an inverse system of BL-algebras. Then (A, φ_i) is an inverse limit of the inverse system $\{A_i, \varphi_{ij}, I\}$, where $A = \{(x_i) \in \prod_{i \in I} A_i \mid \varphi_{ij}(x_i) = x_j \text{ for all } i, j \in I \text{ and } i \ge j\}$ and $\varphi_i : A \longrightarrow A_i$ is the restriction of the natural projection $\pi_i : \prod_{i \in I} A_i \longrightarrow A_i$.

Proof. Let $(x_i), (y_i) \in A$. Then $\varphi_{ij}(x_i) = x_j$ and $\varphi_{ij}(y_i) = y_j$, whenever $i \ge j$. We have $(x_i) \odot (y_i) = (x_i \odot y_i)$ and $(x_i) \to (y_i) = (x_i \to y_i)$, also

$$\varphi_{ij}(x_i \odot y_i) = \varphi_{ij}(x_i) \odot \varphi_{ij}(y_i) = x_j \odot y_j$$

and

$$\varphi_{ij}(x_i \to y_i) = \varphi_{ij}(x_i) \to \varphi_{ij}(y_i) = x_j \to y_j.$$

Therefore $(x_i) \odot (y_i) \in A$ and $(x_i) \to (y_i) \in A$. Thus A is a subalgebra of $\prod_{i \in I} A_i$.

Let $(x_i) \in A$.Then $\varphi_{ij}\varphi_i((x_i)) = \varphi_{ij}(\varphi_i(x_i)) = \varphi_{ij}(x_i) = x_j = \varphi_j((x_i))$, that is φ_i is compatible for all $i \in I$. □

Let B be a *BL*-algebra and $\{\psi_i : B \longrightarrow A_i\}$ be a set of compatible *BL*-homomorphisms. We define $\psi : B \longrightarrow A$ by $\psi(y) = (\psi_i(y))$ for each $y \in B$. The first we show that $\psi(y) \in A$. Since ψ_i's are compatible, so $\varphi_{ij}(\psi_i(y)) = \psi_j(y) \in A_j$, thus $\psi(y) = (\psi_i(y)) \in A$. Now, we will show that ψ is unique. Let $\varphi : B \longrightarrow A$ is another *BL*-homomorphism such that $\varphi_i\varphi = \psi_i$. Also suppose that there exists $y \in B$. such that $\psi(y) \ne \varphi(y)$. By the definition of A, there exists $i \in I$ such that $\varphi_i(\psi(y)) = \psi_i(y) \ne \varphi_i(\varphi(y))$. This is a contradiction. Hence, $\psi(y) = \varphi(y)$ for all $y \in B$. So ψ is unique. Therefore (A, φ_i) is an inverse limit of the inverse system $\{A_i, \varphi_{ij}, I\}$. □

Lemma 3.5. Let $\{A_i, \varphi_{ij}, I\}$ be an inverse system of *BL*-algebras and (A, φ_i) is an inverse limit of $\{A_i, \varphi_{ij}, I\}$. Then the homomorphism $id_A : A \longrightarrow A$ satisfies $\varphi_i id_A = \varphi_i$ for all $i \in I$ and is the only homomorphism with this property.

Proof. Let (A, φ_i) be an inverse limit of *BL*-algebras $\{A_i, \varphi_{ij}, I\}$. By definition (A, φ_i), the maps $\varphi_i : A \longrightarrow A_i$ are compatible. So, the universal property of the inverse limit shows that there exists a unique *BL*-homomorphism $\varphi : A \longrightarrow A$ such that $\varphi_i\varphi = \varphi_i$ for all $i \in I$. The otherhand, since $\varphi_i id_A = \varphi_i$ for all $i \in I$, and φ is unique, $\varphi = id_A$. □

Theorem 3.6. Let (A, φ_i) and (B, ψ_i) are two inverse limits of the inverse system $\{A_i, \varphi_{ij}, I\}$. Then there exists a unique isomorphism $\varphi : A \longrightarrow B$ such that $\psi_i\varphi = \varphi_i$ for all $i \in I$.

Proof. Since the maps $\psi_i : B \longrightarrow A_i$ are compatible, the universal property of the inverse limit (A, φ_i) shows that there exists a unique *BL*-homomorphism $\psi : B \longrightarrow A$ such that $\varphi_i\psi = \psi_i$ for all $i \in I$. Similarly, there exists a unique *BL*-homomorphism $\varphi : A \longrightarrow B$ such that $\psi_i\varphi = \varphi_i$ for all $i \in I$. It follows that $\psi_i = \varphi_i\psi = \psi_i\varphi\psi$ for all $i \in I$. So, by Lemma 3.5, $\varphi\psi = id_B$. Similarly, $\psi\varphi = id_A$. Therefore, φ is an isomorphism. □

Proposition 3.7. Let $\{A_i, \varphi_{ij}, I\}$ be an inverse system of *BL*-algebras and $A = (A, \varphi_i)$ is an inverse limit of $\{A_i, \varphi_{ij}, I\}$. Suppose that $i_0 \in I$ such that $i_0 \ge i_1, \ldots, i_t$ and $\varphi_{i_0 i_k}(X_{i_0}) \subseteq A_{i_k}$ where $X_{i_k} \subseteq A_{i_k}$ for all $k = 0,1,\ldots,t$. Then

A. $A \cap [(\prod_{i \ne i_0} A_i) \times X_{i_0}] = A \cap [(\prod_{i \ne i_0, i_1, \ldots, i_t} A_i) \times X_{i_0} \times \cdots \times X_{i_t}]$.

Proof. Let $(x_i) \in A \cap [(\prod_{i \ne i_0} A_i) \times X_{i_0}]$. Then $(x_i) \in A$ and $(x_i) \in (\prod_{i \ne i_0} A_i) \times X_{i_0}$. Thus, $\varphi_{ij}(x_i) = x_j$ for all $i \ge j$ and $x_{i_0} \in X_{i_0}$. Hence, $\varphi_{i_0 j}(x_{i_0}) = x_j$ for all $j \le i_0$.

Since $i_0 \ge i_1, \ldots, i_t$, $\varphi_{i_0 i_k}(x_{i_0}) = x_{i_k}$ for all $k = 1, \ldots, t$. Since $\varphi_{i_0 i_k}(X_{i_0}) \subseteq X_{i_k}$ for all $k = 1, \ldots, t$, we have $x_{i_k} \in X_{i_k}$. for all $k = 1, \ldots, t$. This implies that

$$(x_i) \in \Big(\prod_{i \ne i_0, i_1, \ldots, i_t} A_i\Big) \times X_{i_0} \times \cdots \times X_{i_t}.$$

Since $(x_i) \in A$, we get

$$(x_i) \in A \cap [(\prod_{i \neq i_0, i_1, \ldots, i_t} A_i) \times X_{i_0} \times \cdots \times X_{i_t}].$$

The other inclusion is easy to show. □

Corollary 3.8. Let $\{A_i, \varphi_{ij}, I\}$ be an inverse system of *BL*-algebras and $A = (A, \varphi_i)$ be an inverse limit of $\{A_i, \varphi_{ij}, I\}$. Then

$$A \cap \left[(\prod_{i \neq i_0} A_i) \times \{0_{A_{i_0}}\}\right] = A \cap \left[\left(\prod_{i \neq i_0, i_1, \ldots, i_t} A_i\right) \times \{0_{A_{i_0}}\} \times \ldots \times \{0_{A_{i_t}}\}\right],$$

where $i_0 \geq i_1, \ldots, i_t$.

Proof. Since $\varphi_{i_0 i_k}$ is a BL-homomorphism for all $i_0 \geq i_1, \ldots, i_t$ and for all $k = 1, \ldots, t$, by Theorem 3.7, $\varphi_{i_0 i_k}(0_{A_{i_0}}) = 0_{A_{i_k}}$ for all $k = 1, \ldots, t$. Thus, $\varphi_{i_0 i_k}(\{0_{A_{i_0}}\}) = \{0_{A_{i_k}}\}$ for all $i_0 \geq i_1, \ldots, i_t$. and for all $k = 1, \ldots, t$. Then by property of *BL*-homomorphism, the result holds. □

Theorem 3.9. Let $\{A_i, \varphi_{ij}, I\}$ be an inverse system of *BL*-algebras and let $A = (A, \varphi_i)$ be its corresponding inverse limit. where $\varphi_i : A \longrightarrow A_i$ is the restriction of the natural projection $\pi_i : \prod_{i \in I} A_i \longrightarrow A_i$. Then for all $j \in I$,

$$A \cap \left[\left(\prod_{i \neq j} A_i\right) \times \{0_{A_j}\}\right] = ker\varphi_j.$$

Proof. For any $j \in I$, we have

$$\begin{aligned} Ker\varphi_j &= \{(x_i) \in A \mid \varphi_j((x_i)) = \pi_j((x_i)) = x_j = 0_{A_j}\} \\ &= \{(x_i) \in A \mid x_j = 0_{A_j}\} \\ &= A \cap \left[(\prod_{i \neq j} A_i) \times \{0_{A_j}\}\right]. \end{aligned}$$

□

Proposition 3.10. Let A be a *BL*-algebra and I be the set of congruences of A. Then $(I, \leq)$ is a directed poset where the binary operation *"≤"* on I is defined by $\theta \leq \varphi$ whenever $\varphi \subseteq \theta$ for all $\theta, \varphi \in I$.

Theorem 3.11. Let I be the set of congruences of *BL*-algebra A. Then $\{A/\phi, \varphi_{\phi\theta}, I\}$ is an inverse system where $\theta \leq \varphi$ whenever $\varphi \subseteq \theta$ for all $\theta, \varphi \in I$. And $\varphi_{\phi\theta} : A/\phi \longrightarrow A/\theta$ is the *BL*-homomorphism defined by $\varphi_{\phi\theta}([x]_\phi) = [x]_\theta$.

Proof. By Proposition 3.10, $(I, \leq)$ is a directed poset. Let

$$\begin{aligned} \varphi_{\theta\lambda}\varphi_{\phi\theta}([x]_\phi) &= \varphi_{\theta\lambda}(\varphi_{\phi\theta}([x]_\phi)) \\ &= \varphi_{\theta\lambda}([x]_\theta) \\ &= [x]_\lambda = \varphi_{\phi\lambda}([x]_\phi). \end{aligned}$$

Therefore, $\varphi_{\theta\lambda}\varphi_{\phi\theta} = \varphi_{\phi\lambda}$ whenever $\phi \geq \theta \geq \lambda$. Note that $\varphi_{\theta\theta} : A/\theta \longrightarrow A/\theta$ is defined by $\varphi_{\theta\theta}([x]_\theta) = [x]_\theta$. Thus, $\varphi_{\theta\theta} = id_{A/\theta}$. Therefore $\{A/\phi, \varphi_{\phi\theta}, I\}$ is an inverse system. □

Theorem 3.12. Let I be the set of congruences of *BL*-algebra A. Consider the inverse system $\{A/\phi, \varphi_{\phi\theta}, I\}$. Then

$$\hat{A} = \left\{([x]_\phi) \in \prod_{\phi \in I} \frac{A}{\phi} \,\middle|\, \varphi_{\phi\theta}([x]_\phi) = [x]_\theta \text{ for all } \phi, \theta \in I, \phi \geq \theta\right\}$$

together with the projections $\varphi_\phi : A \longrightarrow A/\phi$ for all $\phi \in I$ is an inverse limit of $\{A/\phi, \varphi_{\phi\theta}, I\}$.

Proof. The result follows from Theorem 3.4. □

Lemma 3.13. Let A be a *BL*-algebra and $\{A/\phi, \varphi_{\phi\theta}, I\}$ be the defining inverse system of the completion $\hat{A}$ of A. Then the canonical epimorphisms $\varphi_\theta : A \longrightarrow A/\theta$ are compatible where $\theta \in I$.

Proof. Let $x \in A$. Then

$$\varphi_{\phi\theta}\left(\varphi_\phi(x)\right) = \varphi_{\phi\theta}([x]_\phi) = [x]_\theta = \varphi_\theta(x)$$

whenever $\phi \geq \theta$. Therefore, the canonical epimorphisms φ_θ's are compatible. □

Theorem 3.14. Let A be a *BL*-algebra and $\{A/\phi, \varphi_{\phi\theta}, I\}$ be the defining inverse system of the completion $\hat{A}$ of A. Then the canonical epimorphisms $\varphi_\theta : A \longrightarrow A/\theta$ induce a homomorphism $\gamma : A \longrightarrow \hat{A}$ defined by $\gamma(x) = (\varphi_\theta(x)) = ([x]_\theta)$.

Proof. By Lemma 3.13, $\{\varphi_\theta : A \longrightarrow A/\theta\}$ is a set of compatible *BL*-homomorphism. Since A is a *BL*-algebra and $\hat{A} = \lim_{\leftarrow \theta \in I} A/\theta$, by Definition 3.3 and from the proof of Theorem 3.4, there exists a homomorphism $\gamma : A \longrightarrow \hat{A}$ with $\gamma(x) = (\varphi_\theta(x)) = ([x]_\theta)$ for each $x \in A$. □

Theorem 3.15. Let A be a *BL*-algebra and $\{A/\phi, \varphi_{\phi\theta}, I\}$ be the defining inverse system of the completion $\hat{A}$ of A. Then $\gamma^{-1}(([x]_\theta)) = \cap_{\theta \in I} [x]_\theta, for\ all\ ([x]_\theta) \in \hat{A}$.

Proof. Let $y \in \gamma^{-1}(([x]_\theta))$. *Then* $\gamma(y) = ([x]_\theta)$. By the definition of γ, $\gamma(y) = ([y]_\theta)$. Thus, $([y]_\theta) = ([x]_\theta)$. Hence, $[y]_\theta = [x]_\theta$ for all $\theta \in I$. This implies that $y \in \cap_{\theta \in I} [x]_\theta$. It follows that $\gamma^{-1}(([x]_\theta)) \subseteq \cap_{\theta \in I} [x]_\theta$.

Let $y \in \cap_{\theta \in I} [x]_\theta$. Then $y \in [x]_\theta$ for all $\theta \in I$. Thus, $[y]_\theta = [x]_\theta$ for all $\theta \in I$. Hence, $([y]_\theta) = ([x]_\theta)$. Since $\gamma(y) = ([y]_\theta)$, it follows that $\gamma(y) = ([x]_\theta)$. Hence, $y \in \gamma^{-1}(([x]_\theta))$. Therefore $\cap_{\theta \in I} [x]_\theta \subseteq \gamma^{-1}(([x]_\theta))$. □

CONCLUSION

BL-algebras are the algebraic structures for Basic Logic (*BL*, for short), arising from the continuous triangular norms (t-norm), familier in the frameworks of fuzzy set theory.

An inverse limit (also called the projective limit) is a construction that allows one to "glue together" several related objects, the precise gluing process being specified by morphisms between the objects. Inverse limits are a special case of the concept of limit in category theory. In this note, we introduced the notion of inverse limit of an inverse system of *BL*-algebras. Also, we considered the set of congruences of a *BL*-algebra as the poset, then we constructed a quotient inverse system and quotient inverse limit on it.

ACKNOWLEDGMENT

We would like to thank the reviewers for their thoughtful comments and efforts towards improving our manuscript.

REFERENCES

[1] J. Ad´amek, H. Herrlich and G. E. Strecker, Abstract and Concrete Categories (John Wiley & Sons, Inc., 1990). Vol. 17, Springer-Verlag, 1990.

[2] J. Albaracin, J. Vilela, Inverse Limit of an Inverse System of *BE*-algebras, Eur. J. Pure Appl. Math, Vol. 14, No. 2, 2021, pp. 423-430.

[3] S. Eilenberg and N. Steenrod, Foundations of Algebraic Topology, Princeton, New Jersey, Princeton University Press, 1952.

[4] P. H'ajek, Metamathematics of Fuzzy Logic, Kluwer Academic Publishers, Dordrecht, 1998.

[5] N. C. Phillips, Inverse Limits of *C**-Algebras, Journal of Operator Theory, Vol. 19, No. 1,1998, pp. 159-195.

[6] B. Zhao, The Inverse Limit in the Category of Topological Molecular Lattices, Fuzzy Sets and Systems, Vol. 118, Is. 3, 2001, pp. 547-554.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Special types of NM-algebras

1st Saeide Zahiri, 2nd Arsham Borumand Saeid

1st Department of Mathematics, Faculty of Science, *Higher Education Center of Eghlid*, Eghlid, Iran,
saeede.zahiri@yahoo.com, s.zahiri@eghlid.ac.ir
2nd Department of Pure Mathematics, Faculty of Mathematics and Computer,, *Shahid Bahonar University of Kerman,* , Kerman, Iran,
arsham@uk.ac.ir

***Abstract*— The main goal of this paper is to investigate (strong) NM-algebras and study some properties of them. Firstly, we define the notion of perfect NM-algebras and perfect filters on NM-algebras and give some theorems that characterize these algebraic structures. Therefore, the relations between proper NM-algebras and some filters, such as perfect filters and maximal filters have been established. Additionally, we introduce the notion of local NM-algebras and propose the theorems that characterize these algebraic structures.**

***Index Terms*—NM-algebra, Filter, Perfect NM-algebra, Local NM-algebra.**

I. Introduction

In past several years, the study of fuzzy logic and fuzzy reasoning has greatly increased. Some scholars obtained many interesting results on t-norm based fuzzy logics. Hájek's basic logic system BL is an ideal formalization or logical framework of continuous t-norms and their residua. Based on Hájek's work, Esteva and Godo proposed a new formal deductive system MTL, called monoidal t-norm based logic, intended to cope with left-continuous t-norms and their residua. As important schematic extensions of MTL, they gave weak nilpotent minimum logic WNM, the involutive monoidal t-norm based logic IMTL and nilpotent minimum logic NM. The system NM is a common schematic extension of WNM and IMTL which is an ideal formalization of Fodor's nilpotent minimums. We know that nilpotent minimums are important t-norms and possess many excellent properties which are proposed by Fodor in 1995 [7].

The theory of NM-algebras has been enriched with structure theorems [1]. Many of these results have a strong impact with fuzzy logic. In particular, from the structure theory of NM-algebras, one established stone topological representation theorem and characterize the irreducibly representation theorem for NM-algebras. NM-algebras play an important role in studying fuzzy logics and the related algebraic structures. The filter theory of the NM-algebras plays an important role in studying these algebras and the completeness of the NM-logic.

Perfect MV-algebras do not contain non-simple subdirectly irreducible MV-algebras. In fact, the MV-algebra C, introduced by Chang. The MV-algebra C is as a result a proto typical perfect MV-algebra, and we shall see it play a crucial role in the hypothesis of perfect MV-algebras, as well as in the variety it generates [2].

L. P. Belluce et al. introduced the notion of local MV-algebras, while E. Turunen et al. studied the concept of local BL-algebras [6]. By definition, BL-algebras become local if they have a unique maximal deductive system, thereby generalizing the correspondent concept for MV-algebras.

In the present papre, we define the notion of perfect NM-algebras, by some examples explain them and so give some property of these structur algebraic. We consider the relationship between perfect NM-algebras and their perfect filters, and show that an NM-algebra is perfect iff $L/Rad(L) = \{0, 1\}$.

Therefore, we define local NM-algebras, while proposing further characterizations for these algebras. In addition, we have determined the correlations between local NM-algebras and perfect NM-algebras.

II. Preliminaries

In this section, we summarize some definitions and results about NM-algebras, which will be used in the remainder of the paper.

Definition 1. *[4, 5] An algebraic structure* $(L, \wedge, \vee, *, \rightarrow, 0, 1)$ *of type* $(2, 2, 2, 2, 0, 0)$ *is called an* NM*-algebra if it satisfies the following conditions:*

(1) $(L, \vee, \wedge, 0, 1)$ *is a bounded lattice,*
(2) $(L, *, 1)$ *is a commutative monoid,*
(3) $x * y \leq z$ *iff* $x \leq y \rightarrow z$,
(4) $(x \rightarrow y) \vee (y \rightarrow x) = 1$,
(5) $(x * y \rightarrow 0) \vee (x \wedge y \rightarrow x * y) = 1$,
(6) $(x \rightarrow 0) \rightarrow 0 = x$,
for any $x, y, z \in L$.

Proposition 1. *[4, 5] In any NM-algebra $(L, \wedge, \vee, *, \rightarrow, 0, 1)$, the following properties hold: for all $x, y, z \in L$,*

(1) $x \leq y$ *iff* $x \rightarrow y = 1$,
(2) $x \leq y \rightarrow x$
(3) $\neg x = x \rightarrow 0$,
(4) $x \leq y$ *implies* $y \rightarrow z \leq x \rightarrow z$ *and* $z \rightarrow x \leq z \rightarrow y$.

Definition 2. *[4] A nonempty subset F of NM-algebra L is called a filter of L if it satisfies:*

(F_1) $x, y \in F$ *implies* $x * y \in F$,
(F_2) $x \in F, y \in L$ *and* $x \leq y$ *implies* $y \in F$.

We denote by $F[L]$ be the set of all filers of L. A filter F of L is called a proper filter if $F \leq L$.

Let F be a filter of NM-algebra L. Define the congruence $\equiv_F$ on L by $x \equiv_F y$ iff $x \rightarrow y \in F$ and $y \rightarrow x \in F$. The set of all congruence class is denote by L/F, i.e. $L/F = \{[x] | x \in L\}$, where $[x] = \{x \in L | x \equiv_F y\}$. Define $\rightharpoonup, \sqcup, \sqcap, \star$ on L/F as follows: $[x] \rightharpoonup [y] = [x \rightarrow y]$, $[x] \vee [y] = [x \sqcup y]$, $[x] \wedge [y] = [x \sqcap y]$, $[x] * [y] = [x \star y]$. Therefore, $(L/F, \sqcup, \sqcap, \star, \rightharpoonup, [0], [1])$ is an NM-algebra which is called a quotient NM-algebra of L with respect to F.

Definition 3. *[4] For a nonempty subset X of L, we denote by $[X)$ is the filter generated by $[X)$. Clearly, we have $[X) = \{x \in L | x \geq x_1 * x_2 * \ldots * x_n$, for some n and some $x_i \in X\}$. In particular, the principal filter generated by an element $x \in L$ is $[x) = \{y \in L | x_n \leq y, n \geq 1\}$. If F is a filter of L and $x \in L$, then $[F \cup x_i = \{y \in L | f * x_n \leq y$, for some $f \in F, n \in N)$.*

Definition 4. *[4] The order of $v \in L$, delineate by $ord(v)$, is the smallest $k \in \mathbb{N}$ such that $v^k = 0$. If there is no such k, then $ord(v) = \infty$.*

III. Perfect NM-algebras

In this part of paper we donate some of the principal outcome for the theory of perfect NM-algebras. Firstly, the idea of perfect NM-algebras were defined and so some properties of them were considered.

Definition 5. *A NM-algebra L is said to be perfect if for each $s \in L$, $ord(s) \leq \infty$ iff $ord(\neg s) = \infty$.*

Example 1. *Let $L = \{0, a, 1\}$ be a chain. We define operations $*, \rightarrow$ as follow:*

$*$	0	a	1
0	0	0	0
a	0	0	a
1	0	a	1

$\rightarrow$	0	a	1
0	1	1	1
a	a	1	1
1	0	a	1

Then the structure $(L, \wedge, \vee, *, \rightarrow, 0, 1)$ is NM-algebra. It is clear that L is a perfect NM-algebra.

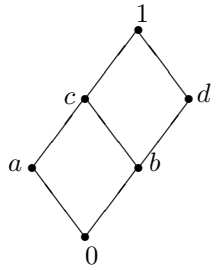
Proposition 2. *Let L_1 and L_2 be perfect NM-algebras. Then $L_1 \times L_2$ is perfect NM-algebra.*

Definition 6. *A filter F in a NM-algerba L is perfect if for each $s \in L$, there exists an n such that $s^n \in F$ iff $\neg(s^m) \notin F$, for any m.*

Example 2. *Let $A = \{0, a, b, c, d, 1\}$. Define*

$*$	0	a	b	c	d	1
0	0	0	0	0	0	0
a	0	a	0	a	0	a
b	0	0	0	0	b	b
c	0	a	0	a	b	c
d	0	0	b	b	d	d
1	0	a	b	c	d	1

$\rightarrow$	0	a	b	c	d	1
0	1	1	1	1	1	1
a	d	1	d	1	d	1
b	c	c	1	1	1	1
c	b	c	d	1	d	1
d	a	a	c	c	1	1
1	0	a	b	c	d	1



Then the structure $(L, \vee, \wedge, *, \rightarrow, 0, 1)$ is an NM-algebra. Clearly, $F = \{1\}$ is perfect filter of L.

Proposition 3. *Let F_1, F_2 be two perfect filters of NM-algebra L. Then $F_1 \cap F_2$ is a perfect filter of L.*

In after theorem we consider the connection between perfect filters and the quotient algebras.

Theorem 1. *A filter F is perfect iff A/F is perfect.*

Proof. Suppose A/F is perfect. If $s \in A$ and $s^k \in F$, for some k, then $ord(\neg s/F) < \infty$. Hence $ord(s/F) = \infty$ and $s^k \neq 0$, for all k and so $\neg(s^k)/F \neq 1/F$, for all k. Thus $\neg(s^k) \notin F$, for all k. If $\neg(s^k) \notin F$, then $ord(s/F) = \infty$ and $ord(\neg s/F) < \infty$ and there is an k such that $s^k \in F$ and $\forall(s^k) \in F$. So F is perfect.

Conversely, suppose F perfect and $ord(s/F) = \infty$. Then $s^k/F \neq 0/F$ and $\neg(s^k)/F \neq 1/F$, for all k. Thus $\neg(s^k) \notin F$, for all k. Since F is perfect we have $\forall s^k \in F$ and so $s^k \in F$, for some k and it follows that $ord(\neg s/F) < \infty$. If $ord(s/F) < \infty$, then $s^k/F = 0$ and $\neg s^k \in F$, for some k. Therefore $\neg(s^k) \notin F$ and s^k/F for all k. So $ord(\neg s/F) = \infty$ and A/F is perfect. □

Definition 7. *A filter F of NM-algebra L is prime, if $x \vee y \in F$, then $x \in F$ or $y \in F$ (or both), for all x and y in L.*

Example 3. *In Example 2, $F = \{a, c, 1\}$ is a prime filter of L.*

Corollary 1. *If the filter F of L is perfect then F is prime.*

In the next theorem the relation between perfect NM-algebras and perfect filters have been considered.

Theorem 2. *NM-algebra L is perfect iff each filter F of L is perfect.*

Proof.

Let all filters of L be perfect. Then filetr $\{1\}$ is perfect. By Theorem 1, $L \equiv L/\{1\}$ is perfect.

Conversely, assume L perfect and $F \subseteq L$ a filter. Let $s \in L$ and $s^k \in F$. Then $ord(s^k) = \infty$ and $ord(\neg(s^k)) < \infty$. So $ord(\neg s) < \infty$. If for some m, $\neg(s^m) \in F$, then $ord(s) < \infty$ which is imposible. Thus there is not m that $\neg(s^m) \in F$. Let $\neg(s^m) \notin F$, for any m. Then $\neg(s^m) \neq 1$ and $s^m \neq 0$, for all m. Therefore $ord(s) = \infty$ and so $ord(\neg s) < \infty$. Hence $\neg(s^k) = 0$, for some k, so $s^k = 1 \in F$. So F is perfect. □

Theorem 3. *In every NM-algebra L, the following hold:*

(i) The class of all perfect NM-algebras is closed under homomorphic images and subalgebras.

(ii) L is perfect iff any proper filter of L is perfect iff $\{1\}$ is a perfect filter.

Definition 8. *A proper filter F of L is maximal filter iff it is not included in any other proper filter of L.*

Example 4. *In Example 2, $M = \{a, c, d, 1\}$ is a maximal filter of L.*

Theorem 4. *Let L be an NM-algebra. M is a maximal filter of L iff for all $x \in L, x \notin M$ there exist $m \in M, n \geq 1$ such that $m * x^n = 0$.*

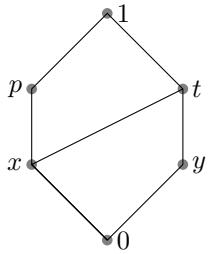
Definition 9. *The intersection of all maximal filter of an NM-algebra L is said to be the radical of L and is shown with $Rad(L)$.*

Example 5. *For $A = \{0, x, y, t, p, 1\}$, define operations $*$ and $\rightarrow$ as follows:*

$*$	0	x	y	t	p	1
0	0	0	0	0	0	0
x	0	0	0	0	0	x
y	0	0	y	y	0	y
t	0	0	y	y	x	t
p	0	0	0	x	p	p
1	0	x	y	t	p	1

$\rightarrow$	0	x	y	t	p	1
0	1	1	1	1	1	1
x	t	1	t	1	1	1
y	p	p	1	1	p	1
t	x	p	t	1	p	1
p	y	t	y	t	1	1
1	0	x	y	t	p	1

Clearly, $(L, \wedge, \vee, *, \rightarrow, 0, 1)$ is an NM-algebra.



Clearly we have :

$$Rad(L) = \{1\}.$$

Theorem 5. *Let L be an NM–algebra. Then*

$$Rad(L) = \{q \in L \mid q \geq \neg(q^k) \text{ for any } k \in \mathbb{N}\}$$

In the following we want to survey the relationship between perfect NM-algebra L and $Rad(L)$.

Proposition 4. *$[Rad(L))$ is a perfect subalgebra of L.*

Proof. Let $q \in [Rad(L))$ and $q \in Rad(L)$. Then $ord(q) = \infty$, so $\neg(q^k) = 1$ and $ord(\neg q) < \infty$.

If $q \in \neg(Rad(L))$, then $\neg q \in Rad(L)$. Hence $ord(\neg q) = \infty$ and $ord(q) < \infty$. □

Definition 10. *In NM-algebra L, we define $\mathcal{M}(L)$ as follows:*

$$\mathcal{M}(L) = \{\neg s \mid s \in L\}.$$

Example 6. *In Example 2, we have $\mathcal{M}(L) = L$.*

Theorem 6. *An NM-algebra L is perfect iff $\mathcal{M}(L)$ is perfect NM-algebra.*

proof. Let L be a perfect NM-algebra. Then for all $\neg x \in \mathcal{M}(L)$,

$$\text{for some } m,\ (\neg s)^m = 0 \text{ iff for all } k,\ (\neg\neg s)^k > 0,$$

i.e. for any $\neg s \in \mathcal{M}(L)$, $ord(\neg s) = \infty$ iff $ord(\neg\neg s) < \infty$, so $\mathcal{M}(L)$ is a perfect NM-algebra.

Conversely, let $s \in L$ and $ord(s) = \infty$. Then for all k, $s^k > 0$ and $ord(\neg s) = \infty$, so $ord(\neg\neg s) = m < \infty$, whence $(\neg s)^m \leq \neg\neg((\neg s)^m) = 0$. Thus $ord(\neg s) < \infty$.

Similary, if $ord(\neg s) = \infty$, then $ord(s) \leq ord(\neg\neg s) < \infty$. Therefore L is perfect. □

IV. Local NM-algebras

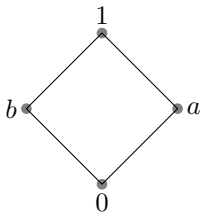
Definition 11. *An NM-algebra L is said to be local iff has exactly one maximal filter.*

Example 7. (i) *In Example 1, it is clear that,* $F = \{1\}$ *is the only maximal filter of* L*. So* L *is local* NM*-algebra.*

(ii) *Let* $L = \{0, a, b, 1\}$*, define* $\odot$ *and* $\Rightarrow$ *on* X *as follows:*

$*$	0	a	b	1
0	0	0	0	0
a	0	a	0	a
b	0	0	b	b
1	0	a	b	1

$\rightarrow$	0	a	b	1
0	1	1	1	1
a	b	1	b	1
b	a	a	1	1
1	0	a	b	1



Then $(L, *, \rightarrow, \vee, \wedge)$ *is an* NM*-algebra. Clearly,* $F_1 = \{a, 1\}$*,* $F_2 = \{b, 1\}$ *are maximal filters of* L*. So* L *is not a local* NM*-algebra.*

Definition 12. *In* NM*-algebra* L*, we define:*

$$D(L) = \{x \in A \mid x^n \neq 0, \textit{ for all } n \in \mathbb{N}\}.$$

Theorem 7. *Let* L *be an* NM*-algebra. Then the following are equivalent:*

(i) $D(L)$ *is a filter,*

(ii) $[D(L))$ *is a proper filter,*

(iii) L *is local,*

(iv) *the unique maximal filter of* A *is* $D(L)$*,*

(v) *if* $x^n, y^n \neq 0$ *for all* $n \geq 1$*, then* $x^n * y^n \neq 0$*, where* $x, y \in L$*.*

proof. $(i \Rightarrow ii)$ Let $D(L)$ be an filter. Then it is easy to see that $[D(L)) = D(L)$ and this filter is proper since $0 \notin D(L)$, so (ii) holds.

$(ii \Rightarrow i)$ if (ii) hold and $x, x \rightarrow y \in D(L) \subseteq [D(L))$, then for all $n \geq 1$ $x^n, (x \rightarrow y)^n \neq 0$ hence $0 \neq x^n * (x \rightarrow y)^n = [x * (x \rightarrow y)]^n \leq y^n$, thus $y \in D(A)$. If $x \in D(L)$, then $x^n \neq 0$, so $\nu x \in D(L)$. Therefore (i) and (ii) are equivalent.

$(i \Leftrightarrow v)$ Since $(x * y)^n = x^n * y^n \leq x^n, y^n$ it is obvious that a necessary and sufficient condition for (i) to hold is (v).

$(iv \Rightarrow iii)$ It is trivial.

$(i \Rightarrow iv)$ Let F' be a filter such that $x \in F', x \notin D(L)$, for some $x \in A$. Then $x^n = 0$, for some $n \in \mathbb{N}$. Hence F' is not proper. So $D(L)$ contains all the proper filters of L and so (iv) holds.

$(iii \Rightarrow iv, i)$ Let L be a local NM-algebra and M_0 be the unique maximal filter of L. Then any element $x \in D(L)$ generates a proper filter $D_x = \{\nu x^n \mid n \geq 0\}$, which can be extended to a maximal filter M_x. But $M_x = M_0$. Thus for all $x \in D(L), x \in M_0$ and so $D(L) \subseteq M_0$. Since M_0 is proper, $M_0 \subseteq D(L)$. Hence $M_0 = D(L)$, therefore (iii) implies (iv) and (i). $\square$

By Theorem 7, we have:

Corollary 2. *Let* L *be a local* NM*-algebra. If* $x^n, y^n \neq 0$ *for all* $n \geq 1$*, where* $x, y \in L$*.*

Proposition 5. *An* NM*-algebra* L *is local iff* $ord(x) < \infty$ *or* $ord(\neg x) < \infty$*, for all* $x \in L$*.*

proof. Let L be a local NM-algebra but $x^n > 0$ and $(\neg x)^n > 0$, for some $x \in L$ and for all $n \in \mathbb{N}$. Then $\neg x, x \in [D(L))$, so $0 = x * x \in [D(L))$, which contradicts to Theorem 7, (ii).

Conversely, let $0 \in [D(L))$. Then for some $x_1, ..., x_n \in D(L)$, we have $x_1 * ... * x_n \leq 0$. Since $ord(x_n) = \infty, ord(\neg x_n) = k_n < \infty$. Thus

$$x_1^{k_n} * ... * x_{n-1}^{k_n} \leq (\neg x_n)^{k_n} = 0,$$

whence

$$x_1^{k_n} * ... * x_{n-2}^{k_n} \leq \neg(x_{n-1}^{k_n}).$$

Clearly, $ord(x_{n-1}^{k_n}) = \infty$. Hence $ord(\neg(x_{n-1}^{k_n})) = k_{n-1} < \infty$, and so

$$x_1^{k_n k_{n-1}} * ... * x_{n-2}^{k_n k_{n-1}} \leq (\neg(x_{n-1}^{k_n}))^{k_{n-1}} = 0.$$

By continuing n times this procedure, we arrive into contradiction $x_1^{k_n ... k_2} = 0$. Therefore $0 \notin [D(A))$, so L is local. $\square$

Corollary 3. *By Definition 5 and Proposition 5 , it is clear that every perfect* NM*-algebra is local.*

Proposition 6. *Let* L *be perfect. Then* $L = [Rad(L))$*.*

proof. Since L is perfect, it is local, so $Rad(L)$ is the unique maximal filter of L and be composed of all elements of finite order. Hence if $q \in L$ and $ord(q) = \infty$, then $q \in Rad(L)$. Let $q \in L$ and $ord(q) < \infty$. Then $ord(\neg q) = \infty$ and $\neg q \in Rad(L)$. Thus $q \in \neg(Rad(L))$ and $L \subseteq [Rad(L))$. $\square$

Remark 1. *If* L *is an* NM*-algebra and* $D \subseteq L$ *is a perfect subalgebra, then* $D = [Rad(D))$*. Now* $Rad(D) = D \cap Rad(L)$ *so we see that* $[Rad(D)) \subseteq [Rad(A))$*.*

Hence $[Rad(A))$ *is a perfect subalgebra of* L *that contains all other perfect subalgebras.*

Definition 13. *Let* M *be a maximal filter of* L*, then we say that* M *is supermaximal if* $L/M = \{0, 1\}$*.*

Corollary 4. L *is perfect iff* $L/Rad(L) = \{0, 1\}$*.*

proof. Let L be perfect. Then $Rad(L)$ is perfect and maximal. So $L/Rad(L)$ is perfect and local, thus $L/Rad(L) = \{0, 1\}$.

Conversely, assume $L/Rad(L) = \{0, 1\}$ chose $q \in L$, then $q/Rad(L) = 0$ or $q/Rad(L) = 1$. So $q \in Rad(L)$ or $q \in \neg Rad(L)$ and $L = [Rad(L))$. □

By Corollary 4, we have:

Corollary 5. *In every perfect NM-algebra L, $Rad(L)$ is supermaximal.*

CONCLUSION

At present, the research papers about perfect NM-algebras have not been proposed. Hence, in order to enrich the theories of algebraic structures, it is meaningful to construct perfect NM-algebras as a generalization of the concept of NM-algebras.

In this paper, motivated by the prior studies on NM-algebras, we prefaced the concept of perfect NM-algebras. We extended and studied these algebras and perfect filters. We have given some characterizations and several examples. Therefore, we used to these results to supply some classification for this algebraic structure. We prove that a filter F on NM-algebra L is perfect iff L/F is perfect.

In addition, we investigated several important properties of local NM-algebras. The special set $D(L)$ was defined, and the correlation between the set and local NM-algebras was determined, while their key properties were also summarized.

In our ulterior work, we will continue our research of algebraic structures, chiefly NM-algebras.

REFERENCES

[1] S. Aguzzoli, B. Gerla, Normal forms and free algebras for some extensions of MTL, Fuzzy Sets and Systems, 159, 2008, pp. 1131-1152.

[2] L. P. Belluce, A. Di Nola, B. Gerla, *Perfect MV-algebras and their logic*, Appl Categor struct 15, 2007, pp. 135-151.

[3] A. Borumand Saeid, S. Zahiri, "Radicals in MTL-algebras", Fuzzy Sets and Systems, 236, 2014, pp. 91-103.

[4] F. Esteva, L. Godo, "Monoidal t-norm based logic: towards a logic for leftcontinuoust-norms", Fuzzy Sets and Systems,124, 2001, pp. 271–288.

[5] M. Bianchi, "First-order Nilpotent minimum logics: first steps", Arch. Math. Logic, 52, 2013, pp. 295–316.

[6] E. Turunen, S. Seesa, "Local BL-algebras", Multiple Valued Logic and soft computing, 6, 2001, pp.229-249.

[7] J. Wang, P. He, Y. She," Monadic NM-algebras", Logic Journal of the IGPL, 27, 2019, pp. 812-835.

[8] S. Zahiri,A. Borumand Saeid,E. Turunen, "On local triangle algebras", Fuzzy Sets and Systems, 418, 2021, pp.126-138.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

The Analysis of A Fractional Network-Based Epidemic Model with Saturated Treatment Function and Fuzzy Transmission

1st Nguyen Phuong Dong, 2nd Hoang Viet Long and 3rd Nguyen Thi Kim Son

1st Dept. Mathematics, *Hanoi Pedagogical University 2*, Vinhphuc, Vietnam. Email: nguyenphuongdong@hpu2.edu.vn
2nd Dept. Information Technology, *People's Police University of Technology and Logistics*, Bacninh, Vietnam. Email: longhv08@gmail.com
3rd Dept. Natural Sciences, *Hanoi Metropolitan University*, Hanoi, Vietnam. Email: ntkson@daihocthudo.edu.vn

***Abstract*— For understanding the influence of malware attacking on complex heterogeneous networks, this work studies a fractional network-based SIRS epidemic model with fuzzy transmission and saturated treatment function. Firstly, we apply the next-generation method to obtain the basic reproductive ratio $\mathcal{R}_0$, that is an important threshold value in the investigation of asymptotic behavior of the proposed epidemic model. The obtained theoretical results indicates that the value $\mathcal{R}_0$ significantly depends on the topology structure of the underlying network and the malware load. In addition, we give a threshold value $\tilde{\mathcal{R}}_0 > \mathcal{R}_0$ that not only determines the existence of endemic equilibrium $\mathbf{E}_*$ but also ensures the clean of malware programs on the network. At last, the sensitivity analysis of the threshold value $\mathcal{R}_0$ and some graphical simulations are presented to illustrate for the theoretical results.**

***Index Terms*—fractional network-based epidemic model, fuzzy transmission, saturated treatment function, basic reproduction number, malware-free equilibrium, endemic equilibrium, asymptotic stability**

I. Introduction

Recently, many researchers have used mathematical modeling based on complex networks to study the spreading of malicious objects in various populations. This approach is known as an effective tool, that helps us to better understand the mechanism of epidemic diseases, to predict the evolution and influence of those diseases on the networks and decide whether they are epidemic or non-epidemic. It is well-known that the nature of epidemic models is the compartmental model, that is, the whole population is divided into some compartments and each compartment contains a number of individuals that share the same epidemiological state. In classical model, when the whole population is small and well-mixed, the rate of disease-causing contacts is often supposed to be equal. This assumption makes the model's evaluation more simply and tractable. However, it is un-realistic when the population is sufficiently large. Indeed, in many kinds of complex networks such as the Internet, Facebook, Instagram social networks, sensor network and biological chain network, etc., the connectability of different nodes on the networks is certainly un-similar and of course, the infections of malware programs to these nodes are also not the same. Therefore, there is a need to taken into consideration the contact heterogeneity of complex networks when mathematically modeling epidemic models of malware program on the networks. Recently, various epidemic models with network-based settings have been analyzed for better understanding the dynamical behavior of epidemic diseases. Indeed, the paper [22] is known as a meaning pioneer work in this topic. This paper presented a study on network-based SIS epidemic model on scale-free network and carried out a detailed study with both analytical and numerical results of the proposed model. The most important contribution is the finding of a threshold value for which the epidemic is absent and the corresponding dynamical behavior. In the paper [8], Huo et. al. proposed a three-compartmental epidemic model with susceptible, infected and recovered states to describe the virus infection on scale-free network. Firstly, the basic reproduction number $\mathcal{R}_0$ was evaluated to study some characteristic properties of the proposed model. After that, by establishing an appropriate Lyapunov function, the authors proved the importance of the number $\mathcal{R}_0$ in the study of asymptotic behavior of endemic equilibrium. Next, the rumor propagation on scale-free network was also studied by Zan et. al. [24], in which the authors formulated an SICR epidemic model and discussed the asymptotic stability of the model's equilibrium points. The contribution of this work is the introduction of a new compartment of counterattack to stifle the rumor propagation. In the paper [13], Li et. al. introduced a SIRS epidemic model to describe the virus propagation on heterogeneous network. This works proved that the presence or absence of the disease on network completely depends on the value of $\mathcal{R}_0$, i.e., the virus-free equilibrium is

globally asymptotically stable if $\mathcal{R}_0 < 1$, while if $\mathcal{R}_0 > 1$ then it is unstable. For better description of the realistic scenario when the number of infected individuals may exceed the treatment capacity, Li and Yousef [12] introduced the saturated treatment function in their work. After formulating a network-based SIRS epidemic model with saturation, the paper [12] calculated the basic reproduction number $\mathcal{R}_0$ and applied it to investigate asymptotic stability of equilibria, the backward bifurcation at $\mathcal{R}_0 = 1$. Moreover, the authors also established a necessary condition for the bifurcation directions at $\mathcal{R}_0 = 1$.

With a long history of development, numerous studies have proved that fractional calculus has a considerable advantages and superiority when modeling many non-local phenomena, the processes with memory and hereditary properties or the motions in viscoelasticity environments. Beside the rapid popularization of fractional calculus, the study of fractional dynamical systems has been paid lots of attentions by researchers and achieved a lot of noticeable results in various fields of basic sciences and engineering. For example, see [3, 4, 6, 7, 10]. In addition, with the introduction of Lyapunov function method for fractional differential systems (see [6]), the stability analysis of fractional differential systems has also attracted a lot of attentions. Due to the better ability in modeling and data-fitting, fractional calculus has been also applied to study the fractional epidemiology theory and applications. However, to the best of our knowledge, there have only a few studies on network-based epidemic models with fractional-order and related problems. Some of them can be found in [5, 7, 10, 21].

Since the nature of almost natural phenomena is vagueness and uncertainty, the mathematical modeling of real-world epidemic diseases must always accept the presence of uncertainties. However, to our best knowledge, there have been very few studies considering the environmental uncertainty in any epidemic model. It is well-known in many biological models that the epidemic disease occurs only if the viral load reaches a certain threshold and obviously, the concept of viral amount is quite difficult to express by exact or certain value. This leads to the use of fuzzy set theory initiated by Zadeh [23] to get the better modeling of epidemic diseases in realistic situations. Despite of the tremendous potential in the modeling of epidemic models, the uses of fuzzy sets in epidemiology theory are not frequent. Some noticeable applications of fuzzy sets in epidemic models can be found in Dong et. al. [3, 4], Mahato et. al. [15], Mondal et. al. [17], Nandi et. al. [19].

Motivated by aforesaid, this work studies a fractional network-based epidemic model with three compartments: Susceptible (S), Infectious (I) and Recovered (R) with fuzzy transmission and the use of saturated treatment. The main contributions of this work can be highlighted as follows:

(i) Based on SIRS epidemic model, we formulate a new epidemic model with fractional-order derivative in the form of mean-field reaction rate equations, namely fractional network-based SIRS epidemic model, for describing and analyzing the spread of malicious objects on scale-free network. Especially, the proposed model considers a fuzzy transmission and a non-linear saturated treatment function for the better fitting with real-world situations.

(ii) Due to the fact that the disease infection often occurs on if the amount of malware program on the network exceeds a certain threshold value and reaches a saturation level at a finite malware load, we propose to use fuzzy membership function to represent the transmission rate.

(iii) Based on the next-generation matrix method, we analytically compute the basic reproduction number $\mathcal{R}_0$, that is an important threshold value in epidemiology theory. This quantity plays a key role in not only the existence of endemic equilibrium $\mathbf{E}_*$ but also the local asymptotic behavior of malware-free equilibrium $\mathbf{E}_0$.

(iv) This work indicates that the proposed epidemic model can't reach the endemic equilibrium state if the basic reproduction number $\tilde{\mathcal{R}}_0 < 1$. In addition, it is also proved that the existence and uniqueness of endemic equilibrium $\mathbf{E}_*$ depends on not only the basic reproduction number $\mathcal{R}_0$ but also the other parameters due to the effect of saturated treatment function.

(v) By using the linearization method and the mathematical induction principle, we give a criteria for the local asymptotic stability or un-stability of malware-free equilibrium $\mathbf{E}_0$ based on the sign of $3n \times 3n$-Jacobi matrix's eigenvalues that are related to the basic reproduction number $\mathcal{R}_0$. Next, by applying the direct Lyapunov functional method with an appropriate Lyapunov function, we can conclude that the attractivity of the equilibrium $\mathbf{E}_0$ depends upon a threshold value $\tilde{\mathcal{R}}_0 > \mathcal{R}_0$, which proves that the condition $\mathcal{R}_0 < 1$ is not sufficient for eliminating the epidemic disease.

The structure of this work is given as follows:

II. Model Formulation

In this paper, we model the complex heterogeneous network by using Barabási-Albert scale-free network [1] to get better description for the heterogeneity of malware spread on complex networks. The structure of Barabási-Albert scale-free network can be briefly summarized as follows: At the initialization, the scale-free network has a small number of fully connected vertices with N_0 nodes and then, a new node with m links is added to the network after each time-step and linked to an old node i with probability $\mathbb{P}(k_i) = \frac{k_i}{\sum_j k_j}$, where k_i is the degree (connectivity) of the i^{th} node. Finally, when the network attains the scale-free stationary state, it can be seen that $\mathbb{P}(k) = ck^{-3}$ is the power-law probability distribution such that a node has k connected links, where c is a parameter such that $\sum_k ck^{-3} = 1$.

A. The fuzzy transmission

In this work, assume that one infectious individual always comes to the contact of maximum one susceptible individual so that the degree-dependent transmission rate of the k^{th}-group $\sigma_k = \sigma k \leq k$. In addition, in order to describe the heterogeneity on the complex network, we propose to represent the transmission rate σ as a function of the available

malware program. In particular, this parameter is proposed to describe through the following fuzzy set:

$$\sigma(\tau) = \begin{cases} 0 & \text{if } \tau \le \tau_m \\ \sigma\frac{\tau-\tau_m}{\tau_0-\tau_m} & \text{if } \tau_m < \tau \le \tau_0 \\ \sigma & \text{if } \tau_0 < \tau \le \tau_M. \end{cases}$$

Here, we can see that there always exists a lower threshold τ_m for the malware propagation, that is, the disease infection occurs only if the amount of malware program on the network must exceed τ_m, otherwise, the chance of transmission is negligible. In addition, the value of τ_m would depend upon both the environmental characteristics and the nature of the malware program, that is reasonable for the choice of fuzzy membership function for transmission rate. Next, there has an upper threshold of the amount of malware program, say τ_0, beyond which the transmission rate reaches the maximum value $\sigma(\tau) = 1$. From τ_m to τ_0, the transmission rate is assumed to vary linearly. Furthermore, assume that the amount of malware program has an upper bound, say τ_M. Moreover, since the nature of realistic phenomena is uncertainty, it is not natural to represent exactly the model's parameters by crisp values. For instance, in order to express the concept "amount of malware program", the use of linguistic variables seems to be more suitable. Indeed, in this work, we assume that the amount of malware programs on the network can be classified into three classes and use linguistic terms, namely "LOW ($\mathcal{A}_l$)", "MEDIUM ($\mathcal{A}_m$)" and "HIGH ($\mathcal{A}_h$)", to characterize for each class. Additionally, in each classification, based on the threshold values τ_m, τ_0, τ_M, the malware load is expressed by using fuzzy numbers (see Figure 1). This approach can be found in [17, 19].

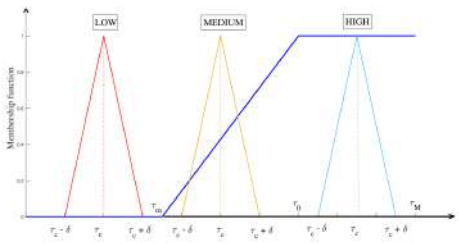


Fig. 1. The membership function of fuzzy transmission rate σ and linguistic variables of the amount of malware program

B. The formulation of the fractional network-based SIRS epidemic model

In the SIR epidemic model, we assume that each node can belong to one of three states: Susceptible state (S), Infectious state (S) and Recovered state (R). In order to taking into consideration the heterogeneity of scale-free networks, the whole population can't be well-mixed and the rate of disease-causing contacts must be varied depending upon the node's connectivity. Indeed, based on the number of connected links a node has per unit time, we classify the whole population into n groups and assume that nodes in a same group are dynamically equivalent. Denote $S_k(t)$, $I_k(t)$, and $R_k(t)$ by the densities of susceptible, infectious and recovered nodes with degree k at time t, respectively for $k = 1, 2, \ldots, n$ and denote $N_k(t)$ by the total number of nodes with degree k at time t. The flowchart of malware propagation of the SIRS epidemic model in the k^{th}-group is given in Figure 2.

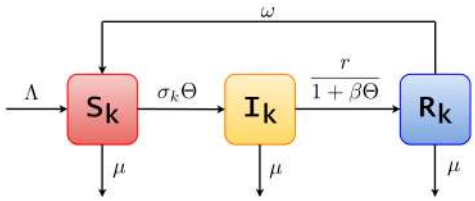


Fig. 2. The flowchart of malware propagation among three compartments: Susceptible (S), Infectious (I), Recovered (R)

In several decades, fractional dynamical systems have proved their importance in real-world modeling due to the effective memory function of fractional derivatives, that has been widely used to model many non-local physical phenomena such as electric flows in signal propagation or processes in the porous media. Motivated by aforesaid, this work is devoted to study a network-based epidemic model governed by the following fractional mean-field reaction rate equation:

$$\begin{cases} {}_0^C\mathfrak{D}_t^\beta S_k(t) = \Lambda - \sigma_k(\tau)S_k(t)\Theta(t) - \mu S_k(t) + \omega R_k(t) \\ {}_0^C\mathfrak{D}_t^\beta I_k(t) = \sigma_k(\tau)S_k(t)\Theta(t) - \mu I_k(t) - \dfrac{rI_k(t)}{1+\gamma\Theta(t)} \\ {}_0^C\mathfrak{D}_t^\beta R_k(t) = \dfrac{rI_k(t)}{1+\gamma\Theta(t)} - (\mu+\omega)R_k(t), \end{cases} \tag{1}$$

with the initial conditions

$$S_k(0) = S_k^0, \quad I_k(0) = I_k^0, \quad R_k(0) = R_k^0, \tag{2}$$

in which the notation ${}_0^C\mathfrak{D}_t^\beta(\cdot)$ denotes for the Caputo fractional derivative of order $\beta \in (0, 1]$ of state functions (see Definition 1 in Appendix). The meanings of the model's parameters are given in Table I:

TABLE I
THE MODEL'S PARAMETERS

No	Parameter	Description
1	$\sigma_k(\tau)$	The degree-dependent fuzzy transmission rate
2	r	The cure rate
3	μ	The natural death rate
4	Λ	The natural birth rate
5	ω	The rate in which a recovered node turns into susceptible

Furthermore, since the un-correlation of node's connectivity on the network is taken into consideration, the probability that a given link is connected to an infectious node can be expressed by the following function

$$\Theta(t) = \frac{1}{\langle k\rangle}\sum_{k=1}^{n} k\mathbb{P}(k)I_k(t),$$

where $\langle k\rangle = \sum_{k=1}^{n} k\mathbb{P}(k)$ is known as the mean degree of the network. On the other hand, since the fact that an anti-malware program only attains a certain maximal treatment capacity for

each epidemic disease, Zhang et. al. [25] introduced a pioneer work on the study of epidemic model with a staged treatment function $h(I) = \frac{rI}{1+\gamma I}$ compatible with the treatment capacity. This treatment function also shows its advantage in measuring the extent of the influence of the infected being delayed for treatment by using a parameter γ in treatment function. This makes our epidemic model seem more reasonable than the case using the linear function. In this paper, the terms $\frac{rI_k}{1+\gamma\Theta}$ represents for the recovery with treatment of the k^{th}-infectious group.

One can easily show that the solution of the fractional differential system (1) with the initial condition (2) is defined for all $t > 0$ and $k = 1, 2, \ldots, n$. From the view point of epidemiology, we only need to focus on the positiveness and the positively invariant set of solution. So we assume that

$$S_k(0) > 0,\ I_k(0) \geq 0,\ R_k(0) \geq 0, \quad k = 1, 2, \ldots, n.$$

Denote

$$\mathbf{x}(t) = (S_1(t), I_1(t), R_1(t), \ldots, S_n(t), I_n(t), R_n(t))^\top$$
$$\Sigma^+ = \left\{ \mathbf{x}(t) \in \mathbb{R}^{3n}_+ : S_k + I_k + R_k \leq \frac{\Lambda}{\mu},\ k = \overline{1, n} \right\}.$$

Due to the presence of epidemic disease on the network and by definition of the probability function $\Theta(t)$, we assume that $\Theta(t) > 0$ for each $t \geq 0$.

Lemma 1. *Assume that*

(i) *The vector function $\mathbf{x}(t)$ is a solution of the fractional network-based SIRS epidemic model (1) with the initial condition (2).*

(ii) *The initial condition (2) belongs to Σ^+.*

Then, for all $t > 0$, the solution $\mathbf{x}(t)$ belongs to Σ^+.

Proof: On the contrary, we assume that for each $k = \overline{1, n}$, there exists a time $t_0 > 0$ such that $S_k(t_0) = 0$, $S_k(t) > 0$ for all $t \in [0, t_0)$ and $S_k(t) < 0$ if $t > t_0$. Then, we consider two following cases:

Case 1: If $I_k(t) \geq 0$ for all $t \geq 0$ then we have

$$\begin{aligned} {}_0^C\mathfrak{D}_t^\beta R_k(t) &= \frac{rI_k(t)}{1+\gamma\Theta(t)} - (\omega+\mu)R_k(t) \\ &\geq -(\omega+\mu)R_k(t). \end{aligned}$$

Then, by applying fractional comparison principle (Lemma 10, [12]), it implies that $R_k(t) \geq R_k(0)\mathbb{E}_\beta\left(-(\omega+\mu)t^\beta\right) \geq 0$ for all $t \geq 0$. As a result, at $t = t_0$, we have

$${}_0^C\mathfrak{D}_t^\beta S_k(t)|_{t=t_0} = \Lambda + \omega R_k(t_0) > 0.$$

By Lemma 4, it yields the function $S_k(t_0) > 0$, which contradicts to our assumption.

Case 2: If there exists a time $t_1 > 0$ such that $I_k(t_1) = 0$, $I_k(t) > 0$ for all $t \in [0, t_1)$ and $I_k(t) < 0$ if $t > t_1$ then our proof is proceeded in two following sub-cases:

Sub-case 1: If $t_1 > t_0$ then by using similar arguments as in Case 1, we can prove that the functions $I_k(t), R_k(t)$ are all non-negative on $[0, t_1]$ and $S_k(t_0) > 0$, which leads to the contradiction.

Sub-case 2: If $t_1 \leq t_0$ then we have $S(t_1) \geq 0$ and

$${}_0^C\mathfrak{D}_t^\beta S_k(t)|_{t=t_1} = \sigma_k(\tau)S_k(t_1)\Theta(t_1) \geq 0.$$

By Lemma 4, we can conclude that $I_k(t)$ is non-increasing on $[0, t_1]$. It should be noted that $I_k(t) > 0$ for all $t \in [0, t_1)$. Thus, it implies that $I_k(t_1) > 0$, that is a contradiction. Therefore, the function $S_k(t) > 0$ for all $t \geq 0$. Similarly, we also get $I_k(t), R_k(t) \geq 0$ for all $t \geq 0$ and $k = \overline{1, n}$.

By the second assumption, we have

$$N_k(0) = S_k(0) + I_k(0) + R_k(0) \leq \frac{\Lambda}{\mu}.$$

By summing up all fractional differential equations of the system (1), we immediately obtain

$${}_0^C\mathfrak{D}_t^\beta N_k(t) = \Lambda - \mu N_k(t). \tag{3}$$

Next, by applying the result of Example 4.9 in [9], the general solution of the fractional differential equation (3) is given by

$$\begin{aligned} N_k(t) &= N_k(0)\mathbb{E}_\beta(-\mu t^\beta) + \Lambda \int_0^t \frac{\mathbb{E}_{\beta,\beta}(-\mu(t-\tau)^\beta)}{(t-\tau)^{1-\beta}} d\tau \\ &= N_k(0)\mathbb{E}_\beta(-\mu t^\beta) + \Lambda t^\beta \mathbb{E}_{\beta,\beta+1}(-\mu t^\beta). \end{aligned}$$

Then, by choosing $\alpha_1 = \beta$, $\alpha_2 = 1$ and $x = -\mu t^\beta$, Lemma 2 implies that

$$\begin{aligned} N_k(t) &= N_k(0)\mathbb{E}_\beta(-\mu t^\beta) + \Lambda t^\beta \mathbb{E}_{\beta,\beta+1}(-\mu t^\beta) \\ &= N_k(0)\mathbb{E}_\beta(-\mu t^\beta) + \frac{\Lambda}{\mu}\left[1 - \mathbb{E}_{\beta,1}(-\mu t^\beta)\right]. \end{aligned}$$

Since $\mathbf{x}(0) \in \Sigma^+$, it implies that $N_k(0) \leq \frac{\Lambda}{\mu}$ and it should be noted that $0 \leq \mathbb{E}_\beta(-\mu t^\beta) \leq 1$ for all $t \geq 0$. Thus, we have

$$N_k(t) \leq \frac{\Lambda}{\mu}\mathbb{E}_\beta(-\mu t^\beta) + \frac{\Lambda}{\mu}\left[1 - \mathbb{E}_\beta(-\mu t^\beta)\right] = \frac{\Lambda}{\mu},$$

which means that Σ^+ is a positively invariant set for the fractional network-based epidemic model (1). ■

III. The basic reproduction number $\mathcal{R}_0$ and equilibrium points

A. The evaluation of basic reproduction number $\mathcal{R}_0$

It can be easily seen that the fractional network-based SIRS epidemic model (1) admits a malware-free equilibrium (MFE) $\mathbf{E}_0 = (\underbrace{\frac{\Lambda}{\mu}, 0, 0, \ldots, \frac{\Lambda}{\mu}, 0, 0}_{3n})$. Now, our aim is to find a threshold value which plays a key role in not only the unique existence of endemic equilibrium $\mathbf{E}_*$ but also the local asymptotic behavior of the model (1). This value is called basic reproduction number and denoted by $\mathcal{R}_0$. For this aim, we propose to apply the next-generation matrix method introduced in [2]. Based on the second equation of the system (1), we can determine the gain terms and lost terms for the epidemic model. Then, the rate matrix $\mathcal{F}$ of new infection's appearance can be given by

$$\mathcal{F} = \frac{\sigma(\tau)\Lambda}{\mu\langle k\rangle} \begin{bmatrix} 1 \\ 2 \\ \vdots \\ n \end{bmatrix} \begin{bmatrix} \mathbb{P}(1) & 2\mathbb{P}(2) & \cdots & n\mathbb{P}(n) \end{bmatrix}$$

and the transition matrix $\mathcal{V}$ of infected states is

$$\mathcal{V} = (r+\mu)\mathtt{I}_n,$$

where $\mathtt{I}_n$ is the $n \times n$ identity matrix. Finally, the basic reproduction number is given by

$$\mathcal{R}_0 = \frac{\sigma(\tau)\Lambda\langle k^2\rangle}{\mu(r+\mu)\langle k\rangle},$$

where $\langle k^2\rangle = \sum_{k=1}^{n} k^2\mathbb{P}(k)$.

Remark 1. *According to the formula of $\mathcal{R}_0$, we can conclude that the threshold value $\mathcal{R}_0$ is directly proportional to the network structure's parameter $\frac{\langle k^2\rangle}{\langle k\rangle}$. This means that the network's heterogeneity can directly affect to the malware widespread on the network.*

B. The existence of an endemic equilibrium

In the following, we discuss the existence of the endemic equilibrium (EE) of the network-based epidemic model (1). For this aim, we denote a threshold value

$$\tilde{\mathcal{R}}_0 = \frac{\Lambda\sigma(\tau)\langle k^2\rangle}{\mu^2\langle k\rangle}.$$

Then, the following theorem presents an interesting result on the existence and uniqueness of an endemic equilibrium of the network-based epidemic model (1).

Theorem 1. *The following assertions are fulfilled:*

1) *If $\tilde{\mathcal{R}}_0 < 1$ then the fractional network-based SIRS epidemic model (1) doesn't have any endemic equilibrium.*
2) *If $\mathcal{R}_0 > 1$ and $\Lambda \leq \mu\left(1+\frac{r}{(\mu+\omega)(1+\gamma)}\right)$ then the fractional network-based SIRS epidemic model (1) has at least one endemic equilibrium $\mathbf{E}_*$*

$$\mathbf{E}_* = (S_1^*, I_1^*, R_1^*, \ldots, S_n^*, I_n^*, R_n^*),$$

where

$$S_k^* = \frac{1}{\sigma_k(\tau)\Theta^*}\left(\mu + \frac{r}{1+\gamma\Theta^*}\right) I_k^*,$$

$$R_k^* = \frac{r}{(\mu+\omega)(1+\gamma\Theta^*)}, \quad \Theta^* = \frac{1}{\langle k\rangle}\sum_{i=1}^{n} i\mathbb{P}(i)I_i^*$$

$$I_k^* = \frac{\Lambda\sigma_k(\tau)\Theta^*}{\mu\left[\mu + \sigma_k(\tau)\Theta^* + \frac{r\sigma_k(\tau)\Theta^*}{(\mu+\omega)(1+\gamma\Theta^*)} + \frac{r}{1+\gamma\Theta^*}\right]}.$$

Moreover, if $\gamma < \frac{\sigma(\tau)}{\mu+\omega}$ then the endemic equilibrium $\mathbf{E}_*$ of the network-based epidemic model (1) is unique.

Proof: Assume that $\mathbf{E}_* = (S_1^*, I_1^*, R_1^*, \ldots, S_n^*, I_n^*, R_n^*)$ is an endemic equilibrium of the fractional network-based SIRS epidemic model (1). Then, for each $k = 1, 2, \ldots, n$, the triple (S_k^*, I_k^*, R_k^*) satisfies the following system

$$\begin{cases} \sigma_k(\tau)S_k\Theta - \mu I_k - \frac{rI_k}{1+\gamma\Theta} &= 0 \\ \frac{rI_k}{1+\gamma\Theta} - (\mu+\omega)R_k &= 0 \\ S_k + I_k + R_k &= \frac{\Lambda}{\mu}, \end{cases} \tag{4}$$

where $\Theta = \frac{1}{\langle k\rangle}\sum_{i=1}^{n} i\mathbb{P}(i)I_i$. Next, by expressing the variables S_k, R_k in the two first equations of the system (4) in the terms of I_k, we immediately get

$$S_k^* = \frac{1}{\sigma_k(\tau)\Theta^*}\left(\mu + \frac{r}{1+\gamma\Theta^*}\right) I_k^*$$

$$R_k^* = \frac{r}{(\mu+\omega)(1+\gamma\Theta^*)} I_k^*.$$

After that, we substitute the expressions of S_k^* and R_k^* into the last equation of the system (4), we receive

$$\left[1 + \frac{1}{\sigma_k(\tau)\Theta^*}\left(\mu + \frac{r}{1+\gamma\Theta^*} + \frac{r}{(\mu+\omega)(1+\gamma\Theta^*)}\right]\right] I_k^* = \frac{\Lambda}{\mu},$$

or equivalently,

$$I_k^* = \frac{\Lambda\sigma_k(\tau)\Theta^*}{\mu\left[\mu + \sigma_k(\tau)\Theta^* + \frac{r\sigma_k(\tau)\Theta^*}{(\mu+\omega)(1+\gamma\Theta^*)} + \frac{r}{1+\gamma\Theta^*}\right]}.$$

Next, by substituting I_k^* into the expression of the function $\Theta(t)$, the equation $\Theta^* = \frac{1}{\langle k\rangle}\sum_{i=1}^{n} i\mathbb{P}(i)I_i^*$ becomes the following self-consistency equation

$$\Theta^* = \frac{1}{\langle k\rangle}\sum_{i=1}^{n} \frac{\Lambda\sigma(\tau)i^2\mathbb{P}(i)\Theta^*}{\mu\left[\mu + \sigma_i(\tau)\Theta^* + \frac{r\sigma_i(\tau)\Theta^*}{(\mu+\omega)(1+\gamma\Theta^*)} + \frac{r}{1+\gamma\Theta^*}\right]}. \tag{5}$$

It should be noted that the self-consistency equation (5) always admits the trivial solution $\Theta \equiv 0$. Now, we aim to determine a sufficient condition for which the equation (5) has a solution $\Theta^* \in (0,1)$. Firstly, we define

$$f(\Theta) = \frac{1}{\langle k\rangle}\sum_{i=1}^{n} \frac{\Lambda\sigma(\tau)i^2\mathbb{P}(i)}{\mu\left[\mu + \sigma_i(\tau)\Theta + \frac{r\sigma_i(\tau)\Theta}{(\mu+\omega)(1+\gamma\Theta)} + \frac{r}{1+\gamma\Theta}\right]}.$$

Here, we can see that

- The function $f(\Theta)$ is continuous on the closed interval $[0,1]$ and differentiable on the open interval $(0,1)$.
- $f(0) = \frac{\Lambda\sigma(\tau)}{\mu(r+\mu)\langle k\rangle}\sum_{k=1}^{n} k^2\mathbb{P}(k) = \mathcal{R}_0$.
- For each $\Theta \in [0,1]$, we have

$$f(\Theta) < \frac{1}{\langle k\rangle}\sum_{i=1}^{n} \frac{\Lambda\sigma(\tau)i^2\mathbb{P}(i)}{\mu^2} = \tilde{\mathcal{R}}_0.$$

- At $\Theta = 1$, we have

$$\begin{aligned} f(1) &= \frac{1}{\langle k\rangle}\sum_{i=1}^{n} \frac{\Lambda\sigma(\tau)i^2\mathbb{P}(i)}{\mu\left[\mu + \frac{r}{1+\gamma} + \sigma_i(\tau)\left(1 + \frac{r}{(\mu+\omega)(1+\gamma)}\right)\right]} \\ &< \frac{1}{\langle k\rangle}\sum_{i=1}^{n} \frac{\Lambda\sigma(\tau)i^2\mathbb{P}(i)}{\mu\sigma_i(\tau)\left(1 + \frac{r}{(\mu+\omega)(1+\gamma)}\right)} \\ &= 1. \end{aligned}$$

Then, the non-trivial solution of the equation (5) is the solution of the following equation

$$\frac{1}{\langle k\rangle}\sum_{i=1}^{n} \frac{\Lambda\sigma(\tau)i^2\mathbb{P}(i)}{\mu\left[\mu + \sigma_i(\tau)\Theta + \frac{r\sigma_i(\tau)\Theta}{(\mu+\omega)(1+\gamma\Theta)} + \frac{r}{1+\gamma\Theta}\right]} = 1. \tag{6}$$

Note that if $\tilde{\mathcal{R}}_0 \leq 1$ then it implies that $f(\Theta) < \tilde{\mathcal{R}}_0 \leq 1$. As a result, there doesn't exist any value $\Theta \in [0,1]$ such that the equation (6) holds, or equivalently, there doesn't exist any endemic equilibrium when $\tilde{\mathcal{R}}_0 \leq 1$. The first assertion of the theorem is completed.

By using the assumption $\mathcal{R}_0 > 1$, it directly follows that $f(0) > 1$. Therefore, by virtue of Intermediate Value theorem, the equation (6) has at least one solution $\Theta \in (0,1)$, that is also the non-trivial solution of the equation (5). As a consequence, the solution $\Theta^* \in (0,1)$ of the self-consistency equation (5) will solve the endemic equilibrium $\mathbf{E}_*$. Therefore, the second assertion holds.

Finally, in order to prove the uniqueness of the endemic equilibrium $\mathbf{E}_*$, let us compute

$$\begin{aligned}\frac{d}{d\Theta}f(\Theta) &= \frac{d}{d\Theta}\left\{\sum_{k=1}^{n}\frac{A_k(1+\gamma\Theta)}{B_k(\Theta)}\right\}\\ &= \sum_{k=1}^{n}\frac{\gamma A_k B_k(\Theta) - A_k(1+\gamma\Theta)\frac{d}{d\Theta}B_k(\Theta)}{B_k^2(\Theta)},\end{aligned}$$

where for simplicity in representation, we denote

$$\begin{aligned}A_k &= \frac{\Lambda\sigma(\tau)k^2\mathbb{P}(k)}{\mu\langle k\rangle},\\ B_k(\Theta) &= (1+\gamma\Theta)(\mu+\sigma_k(\tau)\Theta) + r + \frac{r\sigma_k(\tau)\Theta}{\mu+\omega}.\end{aligned}$$

By some fundamental computations, we obtain

$$\begin{aligned}\frac{d}{d\Theta}f(\Theta) &= \sum_{k=1}^{n}\frac{r\gamma A_k - \frac{r\sigma_k(\tau)A_k}{\mu+\omega} - \sigma_k(\tau)A_k(1+\gamma\Theta)^2}{B_k^2(\Theta)}\\ &= \sum_{k=1}^{n}\frac{rA_k\left(\gamma - \frac{\sigma_k(\tau)}{\mu+\omega}\right)e - \sigma_k(\tau)A_k(1+\gamma\Theta)^2}{B_k^2(\Theta)}.\end{aligned}$$

Therefore, if $\gamma \leq \dfrac{\sigma(\tau)}{\mu+\omega}$ then the derivative $\dfrac{d}{d\Theta}f(\Theta) < 0$ for all $\Theta \in [0,1]$ and hence, the equation (6) has a unique solution $\Theta \in (0,1)$. The proof is completed. ■

IV. The asymptotic behavior of malware-free equilibrium $\mathbf{E}_0$

A. The local asymptotic stability

Theorem 2. *The malware-free equilibrium $\mathbf{E}_0$ of the fractional network-based SIRS epidemic model (1) is locally asymptotically stable if $\mathcal{R}_0 < 1$ and unstable if otherwise.*

Proof: Based on the stability theory of dynamical systems, the local asymptotic stability of the malware-free equilibrium $\mathbf{E}_0$ can be determined by finding the sign of eigenvalues of Jacobi matrix $\mathrm{J}(\mathbf{E}_0)$. Let us consider the Jacobi matrix at $\mathbf{E}_0$ of the epidemic model (1) in the following form

$$\mathrm{J}(\mathbf{E}_0) = \begin{bmatrix}\mathrm{J}_{11} & \mathrm{J}_{12} & \cdots & \mathrm{J}_{1n}\\ \mathrm{J}_{21} & \mathrm{J}_{22} & \cdots & \mathrm{J}_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ \mathrm{J}_{n1} & \mathrm{J}_{n2} & \cdots & \mathrm{J}_{nn}\end{bmatrix}_{3n\times 3n}$$

where J_{kk}, J_{kk} are 3×3-square matrices given by

$$\mathrm{J}_{kk} = \begin{bmatrix}\frac{\sigma_k(\tau)k\Lambda\mathbb{P}(k)}{\langle k\rangle} - (\mu+r) & 0 & 0\\ -\frac{\sigma_k(\tau)k\Lambda\mathbb{P}(k)}{\langle k\rangle} & -\mu & \omega\\ r & 0 & -(\omega+\mu)\end{bmatrix}$$

$$\mathrm{J}_{ki} = \begin{bmatrix}\frac{\sigma_k(\tau)\Lambda i\mathbb{P}(i)}{\langle k\rangle} & 0 & 0\\ -\frac{\sigma_k(\tau)\Lambda i\mathbb{P}(i)}{\langle k\rangle} & 0 & 0\\ 0 & 0 & 0\end{bmatrix} \qquad (k\neq i),$$

for each $k,j = \overline{1,n}$. Then, by applying the mathematical induction principle, the characteristic polynomial w.r.t. $\mathrm{J}(\mathbf{E}_0)$ can be given by

$$\mathcal{P}(\tilde{\lambda}) = (\tilde{\lambda}+\mu)^n(\tilde{\lambda}+\mu+\omega)^n(\tilde{\lambda}+\mu+r)^{n-1}\mathcal{P}_1(\tilde{\lambda})$$

where $\mathcal{P}_1(\tilde{\lambda}) = \tilde{\lambda} + (\mu+r) - \frac{1}{\mu\langle k\rangle}\sum_{k=1}^{n}\sigma_k(\tau)\Lambda k\mathbb{P}(k)$. It can be seen that the characteristic equation $\mathcal{P}(\tilde{\lambda}) = 0$ has three negative solutions: $-\mu$ with multiplicity n, $-(\mu+\omega)$ with multiplicity n and $-(\mu+r)$ with multiplicity $n-1$. As a consequence, the local asymptotic stability of $\mathbf{E}_0$ is only dependent on the sign of the eigenvalue obtained from the equation $\mathcal{P}_1(\tilde{\lambda}) = 0$. Indeed, we have

$$\begin{aligned}\tilde{\lambda} &= -(\mu+r) + \frac{1}{\mu\langle k\rangle}\sum_{k=1}^{n}\sigma_k(\tau)\Lambda k\mathbb{P}(k)\\ &= (\mu+r)\left(\frac{\Lambda\sigma(\tau)}{\mu(\mu+r)\langle k\rangle}\sum_{k=1}^{n}k^2\mathbb{P}(k) - 1\right)\\ &= (\mu+r)(\mathcal{R}_0 - 1).\end{aligned}$$

Therefore, since the assumption $\mathcal{R}_0 < 1$, we can conclude that all eigenvalues of the Jacobi matrix $\mathrm{J}(\mathbf{E}_0)$ are negative and hence, it implies that the malware-free equilibrium $\mathbf{E}_0$ is locally asymptotically stable. Otherwise, we directly imply that this equilibrium is unstable. ■

B. The global asymptotic stability

In the following, we will prove that $\tilde{\mathcal{R}}_0$ is the threshold value for which the malware-free equilibrium $\mathbf{E}_0$ is locally asymptotically stable. Indeed, we have

Theorem 3. *If $\tilde{\mathcal{R}}_0 < 1$ then the malware-free equilibrium $\mathbf{E}_0$ of the fractional network-based SIRS epidemic model (1) is globally asymptotically stable.*

Proof: Assume that $\mathbf{x}(t) = \{(S_k(t), I_k(t), R_k(t))\}_{k=1}^{n}$ is a solution of the fractional network-based SIRS epidemic model (1). Now, we will apply the direct Lyapunov method to discuss the global asymptotic stability of the equilibrium $\mathbf{E}_0$. In particular, we construct the Lyapunov function along $\mathbf{x}(t)$ of by a function $\mathbf{V}: \Sigma^+ \to \mathbb{R}$, given by

$$\mathbf{V}(\mathbf{x}(t)) = \frac{1}{\langle k\rangle}\sum_{k=1}^{n}k\mathbb{P}(k)\left\{\Phi(S_k) + I_k(t) + R_k(t)\right\},$$

where $\Phi(S_k) = S_k(t) - \frac{\Lambda}{\mu} - \frac{\Lambda}{\mu}\ln\left(\frac{\mu S_k(t)}{\Lambda}\right)$. According to Remark A.1, it implies that the function $\mathbf{V}(\mathbf{x}(t))$ is non-

negative definite w.r.t. the malware-free equilibrium $\mathbf{E}_0$. Next, for simplicity in representation, we denote

$$S_k := S_k(t) \qquad I_k := I_k(t)$$
$$R_k := R_k(t) \qquad \Theta := \Theta(t).$$

Next, we take the fractional derivative in Caputo sense for the function $\mathbf{V}(\mathbf{x}(t))$ along $\mathbf{x}(t)$ and apply Lemma 3. One gets

$$\begin{aligned}
{}_0^C\mathfrak{D}_t^\beta \mathbf{V}(\mathbf{x}(t)) &= \frac{1}{\langle k\rangle}\sum_{k=1}^{n} k\mathbb{P}(k)\,{}_0^C\mathfrak{D}_t^\beta \Phi(S_k) \\
&\quad + \frac{1}{\langle k\rangle}\sum_{k=1}^{n} k\mathbb{P}(k)\left({}_0^C\mathfrak{D}_t^\beta I_k + {}_0^C\mathfrak{D}_t^\beta R_k\right) \\
&= \frac{1}{\langle k\rangle}\sum_{k=1}^{n} k\mathbb{P}(k)\left(1-\frac{\Lambda}{\mu S_k}\right){}_0^C\mathfrak{D}_t^\beta S_k \\
&\quad + \frac{1}{\langle k\rangle}\sum_{k=1}^{n} k\mathbb{P}(k)\left({}_0^C\mathfrak{D}_t^\beta I_k + {}_0^C\mathfrak{D}_t^\beta R_k\right),
\end{aligned}$$

where

$$\begin{aligned}
\left(1-\frac{\Lambda}{\mu S_k}\right){}_0^C\mathfrak{D}_t^\beta S_k &= 2\Lambda - \sigma_k(\tau)S_k\Theta - \mu S_k + \omega R_k \\
&\quad - \frac{\Lambda^2}{\mu S_k} + \frac{\sigma_k(\tau)\Lambda\Theta}{\mu} - \frac{\omega\Lambda R_k}{\mu S_k} \\
&= -\frac{\mu}{S_k}\left(\frac{\Lambda^2}{\mu^2} - 2S_k\frac{\Lambda}{\mu} + S_k^2\right. \\
&\quad + \frac{\sigma_k(\tau)\Lambda\Theta}{\mu} - \sigma_k(\tau)S_k\Theta \\
&\quad \left. + \omega R_k\left(1-\frac{\Lambda}{\mu S_k}\right)\right. \qquad (7)
\end{aligned}$$

and

$${}_0^C\mathfrak{D}_t^\beta I_k + {}_0^C\mathfrak{D}_t^\beta R_k = \sigma_k(\tau)S_k\Theta - \mu I_k - (\mu+\omega)R_k. \qquad (8)$$

By combining (7) and (8), we receive

$$\begin{aligned}
&\left(1-\frac{\Lambda}{\mu S_k}\right){}_0^C\mathfrak{D}_t^\beta S_k + {}_0^C\mathfrak{D}_t^\beta I_k + {}_0^C\mathfrak{D}_t^\beta R_k \\
&\le -\frac{\mu}{S_k}\left(\frac{\Lambda}{\mu} - S_k\right)^2 + \frac{\sigma_k(\tau)\Lambda\Theta}{\mu} - \mu I_k \\
&\quad + \omega R_k\left(1 - \frac{\omega+\mu}{\omega} - \frac{\Lambda}{\mu S_k}\right).
\end{aligned}$$

For each $t \ge 0$ and $\mathbf{x}(t) \in \Sigma^+$, note that

$$\begin{aligned}
-\frac{\mu}{S_k}\left(\frac{\Lambda}{\mu} - S_k\right)^2 &\le 0 \\
\omega R_k\left(1 - \frac{\omega+\mu}{\omega} - \frac{\Lambda}{\mu S_k}\right) &\le 0.
\end{aligned}$$

Hence, it follows that

$$\begin{aligned}
{}_0^C\mathfrak{D}_t^\beta \mathbf{V}(\mathbf{x}(t)) &\le \frac{1}{\langle k\rangle}\sum_{k=1}^{n} k\mathbb{P}(k)\left[\frac{\sigma_k(\tau)\Lambda\Theta}{\mu} - \mu I_k\right] \\
&= \frac{\sigma(\tau)}{\mu\langle k\rangle}\sum_{k=1}^{n}\Lambda k^2\mathbb{P}(k)\Theta - \mu\Theta \qquad (9) \\
&= \mu\Theta\left[\frac{\sigma(\tau)}{\mu^2\langle k\rangle}\sum_{k=1}^{n}\Lambda k^2\mathbb{P}(k) - 1\right] \\
&= \mu\Theta(\tilde{\mathcal{R}}_0 - 1).
\end{aligned}$$

Thus, it implies that if $\tilde{\mathcal{R}}_0 < 1$ then ${}_0^C\mathfrak{D}_t^\beta \mathbf{V}(\mathbf{x}(t)) < 0$. In addition, ${}_0^C\mathfrak{D}_t^\beta \mathbf{V}(\mathbf{x}(t)) = 0$ if and only if

$$S_k = \frac{\Lambda}{\mu}, \quad I_k = R_k = 0, \qquad k = 1, 2, \ldots, n.$$

The largest invariant set of $\left\{\mathbf{x}(t) \in \Sigma^+ : {}_0^C\mathfrak{D}_t^\beta \mathbf{V}(\mathbf{x}(t)) = 0\right\}$ is the singleton set $\{\mathbf{E}_0\}$. Therefore, by using Lemma 4.6 in [6], the proof is completed. ■

Remark 2. *By the inequality (9), we have*

$$\begin{aligned}
{}_0^C\mathfrak{D}_t^\beta \mathbf{V}(\mathbf{x}(t)) &\le (\mu+r)\Theta\left[\frac{\Lambda\sigma(\tau)}{\mu(\mu+r)\langle k\rangle}\sum_{k=1}^{n} k^2\mathbb{P}(k) - \frac{\mu}{\mu+r}\right] \\
&\le (\mu+r)\Theta\left(\mathcal{R}_0 - \frac{\mu}{\mu+r}\right).
\end{aligned}$$

This requires $\mathcal{R}_0 \le \dfrac{\mu}{\mu+r} < 1$ *to ensure* ${}_0^C\mathfrak{D}_t^\beta \mathbf{V}(\mathbf{x}(t)) \le 0$. *Therefore, we can conclude that the condition* $\mathcal{R}_0 < 1$ *is not sufficient enough to eliminate the epidemic disease on network. that is the reason why we give a threshold value* $\tilde{\mathcal{R}}_0 > \mathcal{R}_0$ *to evaluate the global asymptotic stability of malware-free equilibrium* $\mathbf{E}_0$.

V. Some discussions

In this section, we formulate the fractional network-based SIRS epidemic model (1) in Barabási-Albert scale-free network, in which we choose $n = 100$ and the probability, that a randomly given node has degree k, is given by $\mathbb{P}(k) = ck^{-3}$. Due to the fact that $\sum_{k=1}^{n}\mathbb{P}(k) = 1$, one gets

$$\sum_{k=1}^{n} ck^{-3} = 1,$$

which follows that the constant $c = 0.8319$. Moreover, the parameters of network structure $\langle k\rangle$ and $\langle k^2\rangle$ are computed by MatLab program as follows:

$$\langle k\rangle = \sum_{k=1}^{n} ck^{-2} \approx 1.3601$$
$$\langle k^2\rangle = \sum_{k=1}^{n} ck^{-1} \approx 4.3154.$$

Moreover, the used parameters of the epidemic model are given in the following table:

TABLE II
The used parameters in the SIRS epidemic model

Parameter	Value	Parameter	Value
Λ	0.12	μ	0.05
σ	0.1	ω	0.06
r	0.8	γ	4

A. The influence of the fuzzy transmission rate to $\mathcal{R}_0$

Since the transmission rate $\sigma_k(\tau) = k\sigma(\tau)$ is represented as a function of malware load τ, the basic reproduction number $\mathcal{R}_0$ then can be known as a fuzzy number w.r.t. the malware load. Based on the analysis results presented in Section III and Section IV, the threshold value $\mathcal{R}_0$ has an essential role in the

asymptotic behavior of the model. In the following, we will discuss the influence of the malware load to the threshold value $\mathcal{R}_0$ and the malware infection. We assume that the amount of malware τ in the population has a linguistic meaning classified as "LOW", "MEDIUM" and "HIGH".

Case I. If the amount of malware program is "LOW", i.e., the triangular fuzzy number $\mathcal{A}_l = (\tau_c - \delta, \tau_c, \tau_c + \delta)$ satisfies $\tau_c + \delta < \tau_m$, then the transmission rate $\sigma_k(\tau) = 0$. In addition, it is clear that the basic reproduction number $\mathcal{R}_0$ then becomes zero, which means that the disease vanishes from the network, i.e., the malware-free equilibrium $\mathbf{E}_0$ is asymptotically stable. This case can be understood that the disease is not enough to cause the infection or malware programs attacked to some nodes that have less importance in the network.

Case II. If the amount of malware program is "MEDIUM", i.e., the triangular fuzzy number $\mathcal{A}_m = (\tau_c - \delta, \tau_c, \tau_c + \delta)$ satisfies $\tau_c - \delta \geq \tau_m$ and $\tau_c + \delta < \tau_0$, then the transmission rate $\sigma_k(\tau)$ is considered a linear function w.r.t. the malware load τ. As a consequence, we also deduce that the basic reproduction number $\mathcal{R}_0 := \mathcal{R}_0(\tau)$, given by

$$\mathcal{R}_0(\tau) = \frac{\Lambda\sigma\langle k^2\rangle}{\mu(r+\mu)\langle k\rangle}\frac{\tau - \tau_m}{\tau_0 - \tau_m},$$

is an increasing function w.r.t. the malware load τ. It leads to a fact that the higher amount of malware program is, the bigger value the basic reproduction number $\mathcal{R}_0$ gets.

Case III. If the amount of malware program is "HIGH", i.e., the triangular fuzzy number $\mathcal{A}_h = (\tau_c - \delta, \tau_c, \tau_c + \delta)$ satisfies $\tau_c - \delta \leq \tau_0$, then the transmission rate $\sigma_k(\tau) = \sigma k$ is a constant function w.r.t the malware load τ. Therefore, the basic reproduction number $\mathcal{R}_0$ only depends on the model's parameters.

B. The sensitivity analysis of the threshold value $\mathcal{R}_0$

Now, we will discuss how different parameters contribute to the change of the threshold value $\mathcal{R}_0$ by evaluating the normalized sensitivity indices. According to Nakul et. al. [18], the sensitivity index of a quantity $\mathbf{x}$ depending on a parameter λ can be determined by $\Upsilon^{\mathbf{x}}_{\lambda} = \dfrac{\partial \mathbf{x}}{\partial \lambda} \times \dfrac{\lambda}{\mathbf{x}}$. By the definition of the basic reproduction number $\mathcal{R}_0$, this quantity depends on some model's parameters such as $r, \sigma(\tau), \mu, \Lambda$ and the parameter of network structure $\frac{\langle k^2\rangle}{\langle k\rangle}$. Therefore, by direct computations, we obtain

$$\Upsilon^{\mathcal{R}_0}_{\sigma(\tau)} = 1 \qquad \Upsilon^{\mathcal{R}_0}_{\Lambda} = 1 \qquad \Upsilon^{\mathcal{R}_0}_{r} = -\frac{r}{\mu + r}$$

$$\Upsilon^{\mathcal{R}_0}_{\mu} = -\frac{(2\mu + r)}{\mu + r} \qquad \Upsilon^{\mathcal{R}_0}_{\frac{\langle k^2\rangle}{\langle k\rangle}} = 1.$$

Remark 3. *We can see that the threshold value $\mathcal{R}_0$ is the most sensitive with the natural death rate μ. Furthermore, we can conclude that the increase of the cure rate r will reduce the value of $\mathcal{R}_0$. In addition, the nodes with different degrees will get different influences to the value $\mathcal{R}_0$. For the fuzzy transmission rate $\sigma(\tau)$, it will experience a 10% increase of the value $\mathcal{R}_0$ if we increase the parameter σ by a same percentage. Similarly, we can also conclude that the value of the basic reproduction number $\mathcal{R}_0$ increases with the increase of the structure parameter $\frac{\langle k^2\rangle}{\langle k\rangle}$, which means that the epidemic disease could be controlled if the value $\frac{\langle k^2\rangle}{\langle k\rangle}$ is decreasing, whereas the higher value of $\frac{\langle k^2\rangle}{\langle k\rangle}$ could follow that more efforts must be done to eliminate malicious objects on the network, i.e. the controlling of epidemic disease becomes more difficult if the parameter $\frac{\langle k^2\rangle}{\langle k\rangle}$ is increasing. The results of sensitive test can be summarized in Table III.*

TABLE III
THE SENSITIVITY INDICES OF MODEL'S PARAMETERS

No	Parameter	Description	Sensitivity index
1	$\sigma(\tau)$	The fuzzy transmission rate	$+1$
2	r	The cure rate	$-\frac{16}{17}$
3	μ	The natural death rate	$-\frac{18}{17}$
4	Λ	The natural birth rate	$+1$
5	$\frac{\langle k^2\rangle}{\langle k\rangle}$	The parameter of network structure	$+1$

In the following, we discuss the change of the basic reproductive number $\mathcal{R}_0$ with respect to malware load τ. Let us choose the normalized values of threshold quantities τ_m, τ_0, τ_M by $0.25, 0.65, 1$, respectively. Then, the transmission rate σ can be represented as a trapezoidal fuzzy number $\tilde{\sigma} = \sigma(0.25, 0.65, 1, 1)$. In the following, we will apply the horizontal membership function approach for fuzzy numbers proposed by Mazandarani et. al. [16] to represent the fuzzy transmission rate $\tilde{\sigma}$ in the parametric form. Indeed, by using this approach, we introduce two parameters α and α_σ, in which $\alpha \in [0, 1]$ is the index of level-sets and $\alpha_\sigma \in [0, 1]$ is the relative-distance-measure variable (see [16]) that measures the granule of information. As a consequence, the horizontal membership function (or gr-representation) of $\tilde{\sigma}$ is

$$\tilde{\sigma}^{gr}(\alpha, \alpha_\sigma) = \sigma\left[0.25 + 0.4\alpha + (0.75 - 0.4\alpha)\alpha_\sigma\right].$$

Then, the relative change of the basic reproduction number $\mathcal{R}_0$ is given in Figure 3.

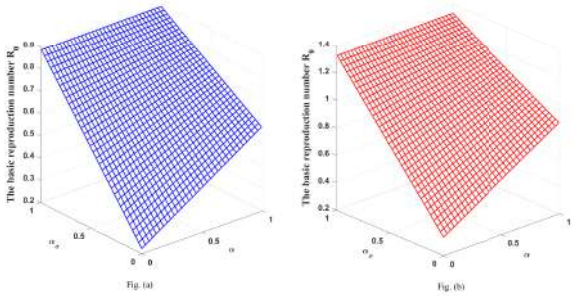


Fig. 3. The relative change of the basic reproduction number $\mathcal{R}_0$ w.r.t. malware load: Fig. (a) $\sigma = 0.1$ and Fig. (b) $\sigma = 0.15$

Figure 3 (b) shows the importance of the malware load in the change of $\mathcal{R}_0$. If the amount of malware programs is increasing then the basic reproduction number $\mathcal{R}_0$ also increases from less than 1 to greater than 1. Hence, there has a noticeable change in the stability state of the proposed

epidemic model when the malware load varies. Additionally, it experienced that the bifurcation occurs at some values τ. This phenomena will be studied in our next work. A similar result was discussed in [19].

VI. Conclusions

This work studied a fractional network-based SIRS epidemic model with fuzzy transmission and saturated treatment function to discuss the malware attacking on the heterogeneous network. In reality, there may occur a scenario that the quantity of infected individuals who need to be treated may exceed the treatment capacity and reach a saturation level. Here, in order to better description for real-world situation, we introduce an epidemic model with a saturated treatment function instead of a linear treatment function. In addition, this work also use linguistic variables and fuzzy membership function to discuss the influence of malware load in the malware infection on the heterogeneous network. Based on the next-generation matrix, we analytically evaluate the basic reproduction number $\mathcal{R}_0$, that is an important threshold value to investigate the asymptotic stability of malware-free equilibrium and the presence of endemic equilibrium on the network. We hope that this work will be the first stage to open up some further studies on the network-based epidemic model such as the condition for the presence of disease, the bifurcation analysis or the quarantine control treatment.

VII. Appendix

Lemma 2. *For each $\beta_1, \beta_2 > 0$, we have*

$$\mathbb{E}_{\alpha_1,\alpha_2}(x) = x\mathbb{E}_{\alpha_1,\alpha_1+\alpha_2}(x) - \frac{1}{\Gamma(\alpha_2)},$$

where $\mathbb{E}_{\alpha_1,\alpha_2}(z)$ is the Mittag-Leffer functions of two parameters α_1 and α_2 (see [9]).

Definition 1. *[9] The Caputo fractional derivative of order $\beta \in (0,1]$ of a function $f \in C^1(a,b)$ is defined by*

$${}_a^C\mathfrak{D}_t^\beta f(t) = \frac{1}{\Gamma(1-\beta)}\int_a^t (t-s)^{-\beta} f'(s)ds.$$

Lemma 3. *[11] Let $\mathbf{x} : [0,\infty) \to \mathbb{R}^+$ be an absolutely continuous function on $[0,\infty)$ and $\beta \in (0,1]$. Then, for each $\mathbf{x}^* \in \mathbb{R}^+$ and $t > 0$, the following inequality holds*

$${}_0^C\mathfrak{D}_t^\beta \left(\mathbf{x}(t) - \mathbf{x}^* - \mathbf{x}^* \ln(\frac{\mathbf{x}(t)}{\mathbf{x}^*})\right) \leq \left(1 - \frac{\mathbf{x}^*}{\mathbf{x}(t)}\right) {}_0^C\mathfrak{D}_t^\beta \mathbf{x}(t).$$

Remark 4. *Let $\Psi : [0,\infty) \to \mathbb{R}$ be a function given by*

$$\Psi(\mathbf{x}) = \mathbf{x} - \mathbf{x}^* - \mathbf{x}^* \ln\left(\frac{\mathbf{x}}{\mathbf{x}^*}\right).$$

Then, it is true that the function $\Psi(\mathbf{x})$ is a non-negative function and attains the global minimum at the point $\mathbf{x} = \mathbf{x}^*$.

Lemma 4. *[20] Assume that $\beta \in (0,1]$ and both the function Φ and its fractional derivative ${}_0^C\mathfrak{D}_t^\beta\Phi$ belong to the space $C[a,b]$. Then we have*

(i) *If ${}_0^C\mathfrak{D}_t^\beta\Phi(t) \geq 0$ then the function $\Phi(t)$ is a non-decreasing function.*

(ii) *If ${}_0^C\mathfrak{D}_t^\beta\Phi(t) \leq 0$ then the function $\Phi(t)$ is a non-increasing function.*

Acknowledgment

Nguyen Phuong Dong was funded by Vingroup JSC and supported by the PhD Scholarship Programme of Vingroup Innovation Foundation (VINIF), Institute of Big Data, code VINIF.2021.TS.009. This work was also funded by Hanoi Pedagogical University 2 Foundation for Sciences and Technology Development via grant number C.2020-SP2-06.

References

[1] A.L. Barabási, R. Albert, "Emergence of scaling in random networks," *Science*, Vol. 286 (5439), pp. 509-512, 1999.

[2] O. Diekmann, J.A.P. Heesterbeek, J.A.J. Metz, "On the definition and the computation of the basic reproduction ratio $\mathcal{R}_0$ in models for infectious diseases in heterogeneous populations, *J. Math. Bio.*, Vol. 28 (4), pp. 365-382, 1990.

[3] N.P. Dong, H.V. Long, A. Khastan, "Optimal control of a fractional order model for granular SEIR epidemic model", *Commun. Nonlinear Sci. Numer. Simulat.*, Vol. 88, 105312, 2020.

[4] N.P. Dong, H.V. Long, N.L. Giang, "The fuzzy fractional SIQR model of computer virus propagation in Wireless Sensor Network using Caputo Atangana-Baleanu derivatives", *Fuzzy Sets Syst.*, 2021. DOI.10.1016/j.fss.2021.04.012.

[5] H. Hu, X. Yuan, L. Huang, C. Huang, "Global dynamics of an SIRS model with demographics and transfer from infectious to susceptible on heterogeneous networks", *Math. Biosci. Eng.*, Vol. 16 (5), pp. 5729-5749, 2019.

[6] J. Huo, H. Zhao, L. Zhu, "The effect of vaccines on backward bifurcation in a fractional order HIV model", *Nonlinear Anal. Real World Appl.*, Vol. 26, pp. 289–305, 2015.

[7] J. Huo, H. Zhao, "Dynamical analysis of a fractional SIR model with birth and death on heterogeneous complex networks", *Physica A*, Vol. 448, pp. 41-56, 2016.

[8] H.F. Huo, P. Yang, H. Xiang, "Dynamics for an SIRS epidemic model with infection age and relapse on a scale-free network", *J. Franklin Inst.*, Vol. 356 (13), pp. 7411-7443, 2019.

[9] A.A. Kilbas, H.M. Srivastava, J.J. Trujillo, Theory and Applications of Fractional Differential Equations, vol. 204 of North-Holland Mathematics Studies, Elsevier, Amsterdam, The Netherlands, 2006.

[10] V.P. Latha, F.A. Rihan, R. Rakkiyappan, G. Velmurugan, "A fractional-order model for Ebola virus infection with delayed immune response on heterogeneous complex networks", *J. Comput. Appl. Math.*, Vol. 339, pp. 134–146, 2018.

[11] C. Vargas-De-León, "Volterra-Type Lyapunov functions for fractional-order epidemic systems," *Commun. Nonlinear Sci. Numer. Simul.*, Vol. 24, pp. 75–85, 2015.

[12] Y. Li, Y. Chen, I. Podlubny, "Mittag-Leffler stability of fractional order nonlinear dynamic systems", *Automatica*, Vol. 45 (8), 1965-1969, 2009.

[13] C.H. Li, C.C. Tsai, S.Y. Yang, “Analysis of epidemic spreading of an SIRS model in complex heterogeneous networks”, *Commun. Nonlinear Sci. Numer. Simulat.*, Vol. 19 (4), pp. 1042-1054, 2014.

[14] C. Li, A.M. Yousef, “Bifurcation analysis of a network-based SIR epidemic model with saturated treatment function,” *Chaos*, Vol. 29 (3), 033129, 2019.

[15] P. Mahato, S. Das, S.K. Mahato, “An epidemic model through information-induced vaccination and treatment under fuzzy impreciseness”, *Model. Earth Syst. Environ.*, 2021. DOI.10.1007/s40808-021-01257-7.

[16] M. Mazandarani, N. Pariz, A.V. Kamyad, “Granular differentiability of fuzzy-number-valued functions”, *IEEE Tran. Fuzzy Syst.*, Vol. 26 (1), pp. 310-323, 2018.

[17] P.K. Mondal, S. Jana, P. Haldar, “Dynamical behavior of an epidemic model in a fuzzy transmission”, *Internat. J. Uncertain. Fuzziness Knowledge-Based Systems*, Vol. 23 (5), pp. 651–665, 2015.

[18] C. Nakul, J.M. Hyman, J.M. Cushing, “Determining important parameters in the spread of malaria through the sensitivity analysis of a mathematical model,” *Bull. Math. Biol.*, Vol. 70, pp. 1272-96, 2008.

[19] S.K. Nandi, S. Jana, M. Manadal, T.K. Kar, “Analysis of a fuzzy epidemic model with saturated treatment and disease transmission”, *Int. J. Biomath.*, Vol. 11 (1), pp. 1–18, 2018.

[20] Z.M. Odibat, N.T. Shawagfeh, “Generalized Taylor’s formula”, *Appl. Math. Comput.*, Vol. 186 (1), pp. 286-293, 2007.

[21] H.A.A. El-Saka, A.A.M. Arafa, M.I. Gouda, “Dynamical analysis of a fractional SIRS model on homogeneous networks”, *Adv. Differ. Equ.*, Vol. 144, 2019.

[22] R. Pastor-Satorras, A. Vespignani, “Epidemic spreading in scale-free networks”, *Phys. Rev. Lett.*, Vol. 86, pp. 3200–3203, 2001.

[23] L.A. Zadeh, “Fuzzy Sets”, *Inf. Contr.*, Vol. 8, pp. 338-353, 1965.

[24] Y. Zan, J. Wua, P. Li, Q. Yu, “SICR rumor spreading model in complex networks:Counterattack and self-resistance”, *Physica A: Stat. Mech. its Appl.*, Vol. 466 (20), pp. 159-170, 2014.

[25] X. Zhang, X. Liu, “Backward bifurcation of an epidemic model with saturated treatment function,” *J. Math. Anal. Appl.*, Vol. 348 (1), pp. 433-443, 2008.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

The Behavior of L-graph Automata

1st Elham Raisi Sarbizhan and 2nd Mohammad Mehdi Zahedi

1st Department of Mathematics, Kerman Graduate University of Advanced Technology,Kerman, Iran, e.raisi@student.kgut.ac.ir
2nd Department of Mathematics, Kerman Graduate University of Advanced Technology, Kerman, Iran, zahedi_mm@kgut.ac.ir

***Abstract*— This paper first introduces the L-graph automaton built on an L-graph using a zero-constraint set. Then, we study the behavior of the corresponding automaton when the L-graph is a path L-graph (cycle, complete, fully bipartite). These L-graph automata have some applications in various fields. One of them is the identification of drugs that have the most similar side effects. These concepts and applications have been illustrated with some examples.**

***Index Terms*—L-graph, L-graph automaton, the behavior of the L-graph automaton**

I. Introduction

Leonhard Euler first coined the term graph. Since then, graph theory has been used for the big problems of life Operations Research. The working group AIM minimum rank-special graphs introduced the concept of a zero-forcing process for bounding the minimum rank of graph G [1]. After that, many authors have extended this notion [5]. For modeling natural events that are uncertain and vague, it was first modeled by Zadeh using a fuzzy set [19]. Therefore, these applications have been considered by many authors. In 1975, Kaufman used the concept of the fuzzy graph to solve real-world problems [6]. Rosenfeld, on the other hand, has many applications for fuzzy graphs in various fields [13]. Wee [17] and Santos [14] introduced fuzzy automata. Fuzzy finite automata have many applications. The residuated lattice was introduced by Morgan Ward and Robert P. Dilworth in 1939 [16]. Many researchers have considered this concept and used it in many different branches of science. Many authors, such as Ciric and his co-authors [4], Qiu [9][10], and Tiwari and his co-authors [15], worked on automata theory based on the residuated lattice. Recently, the new definition of graph based on residuated lattice was presented by Zahedi and Raisi Sarbizhan, etc. [11][12]. Just like fuzzy graphs, there are many applications that Mordeson and his colleagues have recently described in detail in their book [8]. In this paper, the notion of a graph built on a residuated lattice (L-graph) is explained. Moreover, an L-graph automaton is constructed on an L-graph by using zero forcing sets. It also has many applications in various fields. Therefore, we have described a number of these applications in this paper. We have also given some examples and theorems for clarification.

II. Preliminaries

In this section, the basic notions of graph [18][3], residuated lattice [16][2], and L-fuzzy automaton [7] are explained.

Definition 1. *[1]*

- *Color-change rule:*
  If G is a graph with each vertex colored either white or black, u is a black vertex of G, and exactly one neighbor v of u is white, then change the color of v to black.
- *Given a coloring of G, the derived coloring is the result of applying the color-change rule until no more changes are possible.*
- *A zero forcing set for a graph G is a subset of vertices Z such that if initially the vertices in Z are colored black and the remaining vertices are colored white, the derived coloring of G is all black.*
- *$Z(G)$ is the minimum of $|Z|$ over all zero forcing sets $Z \subseteq V(G)$.*

Definition 2. *[16] A residuated lattice is an algebra $L = (L, \wedge, \vee, \otimes, \rightarrow, 0, 1)$ such that*

1) *$L = (L, \wedge, \vee, 0, 1)$ is a lattice (the corresponding order will be denoted by $\leq$) with the smallest element 0 and the greatest element 1,*
2) *$L = (L, \otimes, 1)$ is a commutative monoid (i.e. $\otimes$ is commutative, associative, and $x \otimes 1 = x$ holds),*
3) *$x \otimes y \leq z$ if and only if $x \leq y \rightarrow z$ holds (adjointness condition).*

Proposition 1. *[2] Let $(L, \wedge, \vee, \otimes, \rightarrow, 0, 1)$ be a residuated lattice. Then the following properties hold:*
(R_1) $1 * x = x$, *where* $* \in \{\wedge, \otimes, \rightarrow\}$,
(R_2) $x \otimes 0 = 0, 1' = 0, 0' = 1$,
(R_3) $x \otimes y \leq x \wedge y \leq x, y$, *and* $y \leq (x \rightarrow y)$,
(R_4) $(x \rightarrow y) \otimes x \leq y$,
(R_5) $x \leq y$ *implies* $x * z \leq y * z$, *where* $* \in \wedge, \vee, \otimes$,
(R_6) $z \otimes (x \wedge y) \leq (z \otimes x) \wedge (z \otimes y)$,
(R_7) $x \otimes (y \vee z) = (x \otimes y) \vee (x \otimes z)$.

Definition 3. *[11] $G = (\alpha, \beta)$ is called an L-graph on G^* if $\alpha : V \to L$ and $\beta : E \to L$ are functions, with $\beta(st) \leq \alpha(s) \otimes \alpha(t)$ for every $st \in E$. Besides, if G^* is a path (cycle, bipartite, complete, complete bipartite) graph, then G is called a path (cycle, bipartite, complete, complete bipartite) L-graph on G^*.*

III. The L-graph automaton

In this section, we introduce the notion of a related L-graph automaton construct on the L-graph. In this paper, we assume that L is a residuated lattice and G^* is a simple graph (V, E). It proves the behavior of the associated automaton if the L-graph is a path (cycle, complete, fully bipartite) L-graph.

Definition 4. *Let $G = (\alpha, \beta)$ be an L-graph on G^* and let $Z(G)$ be a zero forcing set of G^*. Then an L-graph automaton $A(Z(G))$ is defined with respect to G is defined by five-tuple (Q, X, μ, F, σ), where;*

(i) *$Q = V$ is the finite nonempty set of states,*
(ii) *$X = \{f, n\}$ is the set of letters of alphabet,*
(iii) *$\mu : V \times X \times V \longrightarrow L$ is the transition function;*
$$\mu(q_i, u, q_j) = \begin{cases} \beta(q_i q_j) & if\ u = f, \\ 1 & if\ u = n, \end{cases}$$
(iv) *F is the set of final states, where $q \in F$ if and only if q does not force any enforce vertices,*
(v) *$\sigma : V \longrightarrow L$ is the initial distribution;*
$$\sigma(q_i) = \begin{cases} 1 & if\ q_i \in Z(G), \\ 0 & otherwise, \end{cases} \quad \text{for every } q_i \in Q.$$

Moreover, a new set $Z(A(Z(G))) = \{q_i \in Q |\ \sigma(q_i) = 1\} = \{q_i \in Q |\ q_i \in Z(G)\}$ has been defined. Also, a response function of $A(Z(G))$ is defined. A k-behavior of $A(Z(G))$ with threshold c is set

$$B_k(A(Z(G)), c) = \{x \in X^* |\ \bigvee_{q \in F} r_\mu(x, q) > c \text{ and } |x| \leq k\}.$$

Definition 5. *Let $A_1 = (Q_1, X, \mu_1, F_1, \sigma_1)$ and $A_2 = (Q_2, X, \mu_2, F_2, \sigma_2)$ be L-graph automata. Then they are equivalent if and only if, $B_k(A_1, c) = B_k(A_2, c)$, for every $k = 0, 1, 2, \ldots$ and for every $c \in L \setminus 1$.*

Example 1. *Suppose $L = ([0,1], \wedge, \vee, \otimes, \to, 0, 1)$, where $a \otimes b = ab$, and $a \to b = \begin{cases} 1 & if\ a \leq b, \\ \frac{b}{a} & if\ a > b, \end{cases}$ and the L-graph $G = (\alpha, \beta)$ on $G^* = (V, E)$, as in Fig. 1, where $V = \{q_1, q_2, q_3, q_4, q_5, q_6\}$, $E = \{q_1q_2, q_1q_6, q_2q_3, q_2q_6, q_3q_4, q_4q_5, q_4q_6, q_5q_6\}$, $\beta(q_iq_j) = \alpha(q_i) \otimes \alpha(q_j)$, $\alpha(q_1) = 0.1$, $\alpha(q_2) = 0.8$, $\alpha(q_3) = 0.6$, $\alpha(q_4) = 0.9$, $\alpha(q_5) = 0.3$, $\alpha(q_6) = 0.5$, $\beta(q_1q_2) = 0.08$, $\beta(q_1q_6) = 0.05$, $\beta(q_2q_3) = 0.48$, $\beta(q_2q_6) = 0.4$, $\beta(q_3q_4) = 0.54$, $\beta(q_4q_6) = 0.45$, $\beta(q_4q_5) = 0.18$, $\beta(q_5q_6) = 0.15$, and $Z(G) = \{q_1, q_2\}$. Then, corresponding L-graph automaton $A(Z(G)) = (Q, X, \mu, F, \sigma)$ can be determined with respect to $Z(G)$, as in Fig. 2, where $Q = V$, $X = \{n, f\}$, $F = \{q_5\}$, $\mu(q_1, n, q_2) = \mu(q_2, n, q_1) = 1$, $\mu(q_4, n, q_6) = \mu(q_6, n, q_4) = 1$, $\mu(q_2, n, q_6) = \mu(q_6, n, q_2) = 1$, $\mu(q_1, f, q_6) = 0.05$, $\mu(q_2, f, q_3) = 0.48$, $\mu(q_3, f, q_4) = 0.54$, $\mu(q_4, f, q_5) = 0.18$,*

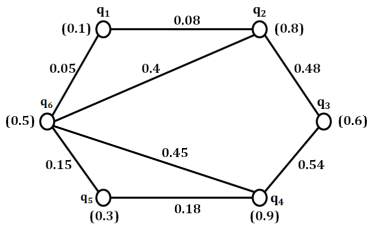


Fig. 1. The L-graph G.

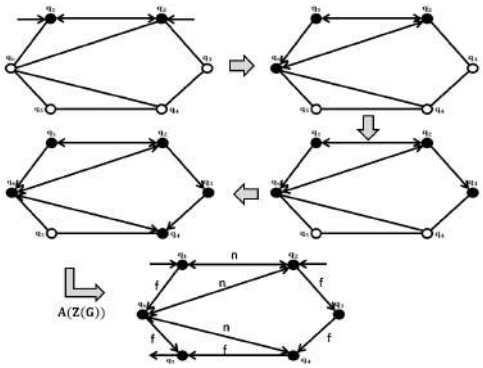


Fig. 2. The L-graph automaton $A(Z(G))$.

$\mu(q_6, f, q_5) = 0.15$, $\sigma(q_1) = \sigma(q_2) = 1$, and $\sigma(q_3) = \sigma(q_4) = \sigma(q_5) = \sigma(q_6) = 0$. So,

$$\begin{aligned} r_\mu(f^3, q) &= \sigma(q_2) \otimes \mu(q_2, f, q_3) \otimes \mu(q_3, f, q_4) \otimes \mu(q_4, f, q_5)) \\ &= 0.48 \otimes 0.54 \otimes 0.18 \\ &= 0.046656. \end{aligned}$$

Also,

$$\begin{aligned} r_\mu(f^2, q) &= \sigma(q_1) \otimes \mu(q_1, f, q_6) \otimes \mu(q_6, f, q_5) \\ &= 0.05 \otimes 0.15 \\ &= 0.075. \end{aligned}$$

Theorem 1. *Let $G = (\alpha, \beta)$ be an L-graph on $G^* = (V, E)$, and let $\beta = \bigvee_{qq' \in E} \beta(qq')$ and $\beta' = \bigwedge_{qq' \in E} \beta(qq')$, Then,*

(i) *If G is a path L-graph with n vertices, then $A(Z(G)) = (Q, X, \mu, F, \sigma)$ is a related L-graph automaton such that*
$$r_\mu(f^{n-1}, q) \leq \otimes_{q \in V} \alpha(q),$$
for every zero forcing set $Z(G)$.
(ii) *If G is a cycle L-graph with $2k$ vertices, then $A(Z(G)) = (Q, X, \mu, F, \sigma)$ is an associated L-graph automaton such that*
$$\otimes_{i=1,\ldots,k-1} \beta' \leq \bigvee_{q \in F} r_\mu(n^* f^{k-1} n^*, q) \leq \otimes_{i=1,\ldots,k-1} \beta,$$
for every zero forcing set $Z(G)$.

(iii) *If G is a cyclic L-graph with $2k+1$ vertices, then $A(Z(G)) = (Q, X, \mu, F, \sigma)$ is an associated L-graph automaton such that*

$$\otimes_{i=1,\ldots,k}\beta' \leq r_\mu(n^* f^k, q) \leq \otimes_{i=1,\ldots,k}\beta,$$

for every zero forcing set $Z(G)$.

(iv) *If G is a complete L-graph with n vertices, then $A(Z(G)) = (Q, X, \mu, F, \sigma)$ is an associated L-graph automaton such that*

$$\beta' \leq r_\mu(f, q) \leq \beta,$$

for any zero forcing set $Z(G)$.

(v) *If G is a complete bipartite L-graph with n vertices, then $A(Z(G)) = (Q, X, \mu, F, \sigma)$ is an associated L-graph automaton such that*

$$\beta' \leq \bigvee_{q\in F} r_\mu(f, q) \leq \beta,$$

for any zero forcing set $Z(G)$.

Proof: (i) G has two distinct zero forcing sets $Z_1(G)$ and $Z_2(G)$, but these related L-graph automata are equivalent and isomorphic. Thus, consider $Z(G) = \{q_1\}$. Thus

$$\begin{aligned} r_\mu(f^{n-1}, q) &= r_\mu(f^{n-1}, q_n) \\ &= \sigma(q_1) \otimes \mu^*(q_1, f^n, q_n) \\ &= \mu(q_1, f, q_2) \otimes \ldots \otimes \mu(q_{n-1}, f, q_n) \\ &\leq \otimes_{q\in V}\alpha(q), \text{ by } Proposition\ 1(R_5). \end{aligned}$$

(ii) Since these zero forcing sets are one of the two vertices adjacent to this L-graph, there are $k-1$ vertices in each chain that constraining the other vertex. Suppose $Z(G) = \{q_1, q_2\}$, and $F = \{k+1, k+2\}$. Thus,

$$\begin{aligned} \bigvee_{q\in F} r_\mu(f^{k-1}, q) &= \sigma(q_1) \otimes \mu^*(q_1, f^{k-1}, q_{k+2}) \\ &\vee \ \sigma(q_2) \otimes \mu^*(q_1, f^{k-1}, q_{k+1}) \\ &= \mu^*(q_1, f^{k-1}, q_{k+2}) \vee \mu^*(q_1, f^{k-1}, q_{k+1}) \\ &\leq \otimes_{i=1,\ldots,k-1}\beta, \end{aligned}$$

and

$$\begin{aligned} \bigvee_{q\in F} r_\mu(f^{k-1}, q) &= \mu^*(q_1, f^{k-1}, q_{k+2}) \vee \mu^*(q_1, f^{k-1}, q_{k+1}) \\ &\geq \otimes_{i=1,\ldots,k-1}\beta'. \end{aligned}$$

Moreover, for any $n^* f^{k-1} n^*$, it is similar to the above with some modifications.

(iii) The proof with some terms is similar to the above.

(iv) We know that every zero forcing set has $n-1$ vertices. $Z(G) = \{q_1, q_2, \ldots, q_{n-1}\}$, and $F = \{q_n\}$. Thus,

$$\begin{aligned} r_\mu(f, q) &= \mu(q_1, f, q_n) \vee \ldots \vee \mu(q_{n-1}, f, q_n) \\ &\leq \beta, \end{aligned}$$

and

$$\begin{aligned} r_\mu(f, q) &= \mu(q_1, f, q_n) \vee \ldots \vee \mu(q_{n-1}, f, q_n) \\ &\geq \beta'. \end{aligned}$$

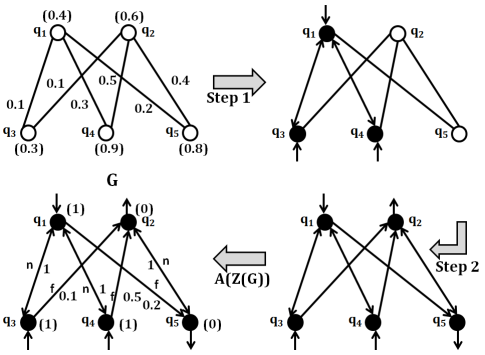


Fig. 3. The L-graph G.

(iv) The proof is similar to the proof above but with some modifications. ∎

Example 2. *Consider L in Example 1 and a complete bipartite L-graph $G = (\alpha, \beta)$ on G^*, as in Fig. 3, where $V = \{q_1, q_2, q_3, q_4, q_5\}$, $E = \{q_1q_3, q_1q_4, q_1q_5, q_2q_3, q_2q_4, q_2q_5\}$, $\alpha(q_1) = 0.4$, $\alpha(q_2) = 0.6$, $\alpha(q_3) = 0.3$, $\alpha(q_4) = 0.9$, $\alpha(q_5) = 0.8$, $\beta(q_1q_3) = 0.1$, $\beta(q_1q_4) = 0.3$, $\beta(q_1q_5) = 0.2$, $\beta(q_2q_3) = 0.1$, $\beta(q_2q_4) = 0.5$, $\beta(q_2q_5) = 0.4$, and $Z(G) = \{q_1, q_3, q_4\}$. Therefore, $A(Z(G)) = (Q, X, \mu, F, \sigma)$ is an associated L-graph automaton, shown in Fig. 3, where $Q = V$, $F = \{q_2, q_5\}$, $\mu(q_1, n, q_3) = \mu(q_3, n, q_1) = 1$, $\mu(q_1, n, q_4) = \mu(q_4, n, q_1) = 1$, $\mu(q_1, f, q_5) = 0.2$, $\mu(q_3, f, q_2) = 0.1$, $\mu(q_4, f, q_2) = 0.5$, $\mu(q_2, n, q_5) = \mu(q_5, n, q_2) = 1$, $\sigma(q_1) = \sigma(q_3) = \sigma(q_4) = 1$, and $\sigma(q_2) = \sigma(q_5) = 0$. So,*

$$\begin{aligned} r_\mu(f, q) &= (\sigma(q_1) \otimes \mu(q_1, f, q_5)) \vee (\sigma(q_3) \otimes \mu(q_3, f, q_2)) \\ &\vee \ (\sigma(q_4) \otimes \mu(q_4, f, q_2)) \\ &= 0.2 \vee 0.1 \vee 0.5. \end{aligned}$$

Also, for every $x = n^* f n^*$, $r_\mu(x, q) = 0.5$.

IV. Some applications of L-graph automata

In this section, we discuss the applications of L-graph automata and illustrate them with an example.

Application 1. *The related L-graph automata have some applications. For example;*

a: *Assume n drugs with m side effects and $L = (P(X), \cap, \cup, \otimes, \to, \emptyset, X)$, where $X = \{a_1, a_2, \ldots, a_m\}$ such that a_i per i is labeled for m side effects of these drugs and $A \otimes B = A \cap B$, and $A \to B = \begin{cases} X & if\ A \subseteq B, \\ B & if\ B \subset A. \end{cases}$ Then, $G = (\alpha, \beta)$ on G^* is the L-graph in which $V = \{q_1, q_2, \ldots, q_n\}$ such that q_i for each $i = 1, 2, \ldots, n$ is labeled with these drugs. Moreover, $q_iq_j \in E$ if and only if these two drugs have at least one similar side effect $\alpha(q_i) = \{a_k | a_k$ is one of side effects of $q_i\}$ and $\beta(q_iq_j) = \alpha(q_i) \otimes \alpha(q_j)$, for each $q_iq_j \in E$.*

Therefore, $A(Z(G)) = (Q, X, \mu, F, \sigma)$ is the associated L-graph automaton such that $Z(G)$ is a zero-forcing set. Moreover, these behaviors whose words do not have the label n are certain drugs that have the most similar side effects.

b: *Suppose that n drugs have m health benefits. As in the above method, we can use $A(Z(G)) = (Q, X, \mu, F, \sigma)$, which is the associated L-graph automaton, so $Z(G)$ is a zero forcing set. Moreover, these behaviors whose words do not have the label n are determined to be drugs that have the most similar health benefits.*

c: *These L-graph automata can also be used to determine the number of articles or books that have the most similar topics.*

Example 3. *Suppose that four drugs q_1, q_2, q_3, and q_4 have five side effects, namely, headache, dizziness, nausea, drowsiness, and physical pain, so that drug q_1 has two side effects: Headache, dizziness, and nausea, and drug q_2 has four side effects: Dizziness, nausea, drowsiness and body pain, the drug q_3 two side effects: Nausea and drowsiness, and drug q_4 one side effect: physical pain. Suppose $L = (P(X), \cap, \cup, \otimes, \rightarrow, \emptyset, X)$, where $X = \{a_1, a_2, \ldots, a_5\}$, $A \otimes B = A \cap B$, and $A \rightarrow B = \begin{cases} X & if\ A \subseteq B, \\ B & if\ B \subset A. \end{cases}$ Then, $G = (\alpha, \beta)$ on G^* is the L-graph, as in Fig. 4, where $V = \{q_1, q_2, q_3, q_4\}$, $E = \{q_1q_2, q_2q_3, q_2q_4\}$, $\alpha(q_1) = \{a_1, a_2, a_3\}$, $\alpha(q_2) = \{a_2, a_3, a_4, a_5\}$, $\alpha(q_3) = \{a_3, a_4\}$, $\alpha(q_4) = \{a_5\}$, $\beta(q_1q_2) = \{a_2, a_3\}$, $\beta(q_2q_3) = \{a_3, a_4\}$, and $\beta(q_2q_4) = \{a_5\}$. For any zero forcing set, these related L-graph automata are equivalent. Consider $Z(G) = \{q_1, q_4\}$. Thus, $A(Z(G)) = (Q, X, \mu, F, \sigma)$ is the related L-graph automaton, as in Fig. 4, where $Q = V$, $\sigma(q_1) = \sigma(q_4) = X$, $\sigma(q_2) = \sigma(q_3) = \emptyset$, $\mu(q_1, f, q_2) = \{a_2, a_3\}$, $\mu(q_4, f, q_2) = \{a_5\}$, and $\mu(q_3, f, q_2) = \{a_3, a_4\}$. Hence,*

$$\begin{aligned} r_\mu(f^2, q) &= (\sigma(q_1) \otimes \mu(q_1, f, q_2) \otimes \mu(q_2, f, q_3)) \\ &\vee (\sigma(q_4) \otimes \mu(q_4, f, q_2) \otimes \mu(q_2, f, q_3)) \\ &= ((\{a_2, a_3\}) \otimes (\{a_3, a_4\})) \\ &\vee ((\{a_5\}) \otimes (\{a_3, a_4\})) \\ &= \{a_5\} \vee \emptyset \\ &= \{a_5\}. \end{aligned}$$

drugs q_1, q_2, and q_3 have the most similar side effects.

V. Conclusion

In this paper, the associated L-graph automata are introduced, and their behaviors are studied. In addition, some theorems and examples are presented for clarification. We also found some applications for L-graph automata. For example: find drugs that have the most similar treatments. L-Graph automata are used to find the drugs that are most similar in terms of side effects or health benefits. One of the applications of this article is to help medical researchers treat diseases better. In the next article, we will try to establish a connection between the L-graph and the corresponding L-graph automaton to solve more complex problems, such as choosing the best drug for a disease.

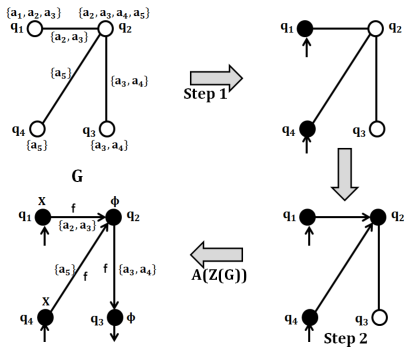


Fig. 4. The L-graph G, and the related L-graph automaton $A(Z(G))$.

References

[1] AIM Minimum Rank–Special Graphs Work Group, "Zero forcing sets and the minimum rank of graphs," Linear Algebra and its Applications, vol. 428, no. 7, p.p 1628-1648, Apr. 2008.

[2] K. Blount, and C. Tsinakis, "The structure of residuated lattices," International Journal of Algebra and Computation, vol. 13, n. 4, pp. 437– 461, Aug. 2003.

[3] J.A Bondy, and U.S. R. Murty, Graph theory, Landon, UK, Springer London, 2008, pp. 29-86.

[4] M. Ciric, "Lattices of subautomata and direct sum decompositions of automata," Algebra Colloquium, vol. 6, no. 1, pp. 71-88, Dec. 1999.

[5] L. Deaett, and S. A. Meyer, "The minimum rank problem for circulants," Linear Algebra and its Applications, vol. 491, no. 4, pp. 386-418, Feb. 2016.

[6] A. Kauffman, Introduction a La theorie des sous-ensembles flous: A L'Usage Des Ingenieurs (Fuzzy Sets Theory), Tome III, Paris, French, Masson Et Cie, Editeurs, 1975.

[7] J.N Mordeson, and D.S Malik, Fuzzy automata and languages theory and applications, Chapman and Hall, 2002, pp. 32-134.

[8] J. N. Mordeson, and S. Mathew, Mathematics of Uncertainty for Coping with World Challenges: Climate Change, World Hunger, Modern Slavery, Coronavirus, Human Trafficking, Springer International Publishing, 2021, pp. 133-173.

[9] D. Qiu, "Automata theory based on complete residuated lattice-valued logic(I)," Sci China Ser F, vol. 44, no. 6 pp. 419–429, Oct. 2001.

[10] D. Qiu, D. "Automata theory based on complete residuated lattice-valued logic (II)," Sci China Ser F, vol. 45, no. 6, pp. 442–452, Oct. 2002.

[11] E. Raisi Sarbizhan, M.M. Zahedi, and M. Sahmsizade, "L-graph automata and some applications," submitted.

[12] E. Raisi Sarbizhan, and M.M. Zahedi, "Maximal product of two L-graph automata and some applications," submitted.

[13] A. Rosenfeld, ”Fuzzy groups,” Journal of Mathematical Analysis and Applications, vol. 35, no. 3, p.p 512-517, Sep. 1971.
[14] E.S Santos, ”Maxmin automata,” Information Control, vol. 13, no. 2, pp.363-377, 1968.
[15] S.p. Tiwari, A.K. Singh, and S. Sharan, ”Fuzzy Automata based on Lattice-Ordered Monoid and Associated Topology,” Journal of Uncertain Systems, vol. 6, no. 1, pp. 51-55, May. 2012.
[16] M. Ward, and R.P.Dilworth, ”Residuated lattices,” Transactions of The American Society, vol. 45, no. 3, pp. 335–354, Apr. 1939.
[17] W. G Wee, On generalization of adaptive algorithm and application of the fuzzy sets conceptto pattern classification, Ph.D. Thesis, Purdue University, Lafayette, India, 1967.
[18] D.B. West, Introduction to graph theory, (Vol. 2), Delhi, India, Pearson Education 2001, pp. 19-67.
[19] L.A. Zadeh, ”Similarity relations and fuzzy orderings,” Information Sciences, vol. 3, no. 2, p.p 177–200, Apr. 1971.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

The characterization of L^B-valued GFA via L^B-valued operators

1st Kh. Abolpour, 2nd M.M. Zahedi, 3rd M. Shamsizadeh

1st Dept. of Math., Shiraz Branch, Islamic Azad University, Shiraz, Iran, Abolpor_kh@yahoo.com
2nd Dept. of Math., Graduate University of Advanced Technology, Kerman, Iran, zahedi_mm@kgut.ac.ir
3rd Dept. of Math., Behbahan Khatam Alanbia University of Technology, Khouzestan, Iran, shamsizadeh.m@gmail.com

***Abstract*— This paper aims to study the L^B-valued general fuzzy automata theory based on t-norm/t-conorm and general implicators and investigate their algebraic and L^B-valued topological properties. Specifically, we associate L^B-valued co-topologies/topologies for a given L^B-valued general fuzzy automaton and use these to characterize some algebraic concepts.**

***Index Terms*— L^B-valued general fuzzy automata, topology, Alexandroff, continuous, operator**

I. Introduction

Since the theory of fuzzy sets was introduced by Zadeh [7], fuzzy automata and languages have been studied as methods for bridging the gap between the precision of computer languages and vaguness. These studies were initiated by Santos [3], Wee [5], and Wee and Fu [6], and further developed by many researchers. Fuzzy automata and languages with membership values in different lattice structures have attracted considerable attention from researchers working in this area. It is well-known that the importance of the algebraic study is due to the fact that the algebraic properties play a vital role in the development of fundamentals of computer science; while, the benefits of (fuzzy) topological methods in the study of the crisp as well as fuzzy automata have been achieved (i) by way of the resulting economy in the arguments used to prove several results, and (ii) the additional insights provided by the established discipline of topology (cf. [2]). The present study is an attempt to investigate the theory of L^B-valued general fuzzy automata, characterizing its properties through some L^B-valued operators based on t-norm/t-conorm and implicators. These L^B-valued operators lead us to characterize some algebraic concepts associated with an L^B-valued general fuzzy automaton and to study L^B-valued topological concepts therein.

II. Preliminaries

In this section, to recall the ideas associated with t-norm, t-conorm, negator, implicators and L^B-valued general fuzzy automata with the following.

Definition 1. *[2] A lattice $(L, \leq, \wedge, \vee, 0, 1)$ is called a **complete lattice** with the greatest element 1 and the least element 0 if every subset (finite as well as infinite) has a supremum and infimum. Throughout this paper, we assume that L is a complete lattice $(L, \leq, \wedge, \vee, 0, 1)$ and for a nonempty set $X, \lambda : X \to L$ is an L-fuzzy set in X. Further, for a nonempty set X, L^X denotes the collection of all L-fuzzy sets in X. Furthermore, for $a \in L, \boldsymbol{a}, \boldsymbol{0}, \boldsymbol{1} : X \to L$ are the maps such that for all $x \in X$, $\boldsymbol{a}(x) = a, \boldsymbol{0}(x) = 0$, and $\boldsymbol{1}(x) = 1$. Moreover, for $x \in X$ and $c \in L$, an L-fuzzy set $x_c : X \to L$ is called an L-fuzzy point if for all $y \in X$,*

$$x_c(y) = \begin{cases} c & if \quad x = y \\ 0 & otherwise. \end{cases}$$

Now, let $f : X \to X'$ be a map. Then according to Zadeh's extension principle, f can be extended to the L-fuzzy operators $f^{\rightarrow} : L^X \to L^{X'}$ and $f^{\leftarrow} : L^{X'} \to L^X$ such that for all $\lambda \in L^X, \lambda' \in L^{X'}$, and $x' \in X'$,

$$f^{\rightarrow}(\lambda)(x') = \vee\{\lambda(x) : x \in X, f(x) = x'\} \text{ and } f^{\leftarrow}(\lambda') = \lambda' \circ f.$$

Definition 2. *[2] A triangular norm (t-norm, for short) $\mathcal{T}$ is a map $\mathcal{T} : L \times L \to L$ such that the following conditions hold:*
(i) $\mathcal{T}(1, a) = a, \forall a \in L$;
(ii) $\mathcal{T}(a, b) = \mathcal{T}(b, a), \forall a, b \in L$;
(iii) $\mathcal{T}(a, \mathcal{T}(b, c)) = \mathcal{T}(\mathcal{T}(a, b), c), \forall a, b, c \in L$;
(iv) if $a \leq c$ and $b \leq d$, then $\mathcal{T}(a, b) \leq \mathcal{T}(c, d)$.
In addition, $\mathcal{T}$ is called a left (respectively, right) continuous if for all $a, b \in L$ and $\{a_j : j \in J\}, \{b_j : j \in J\} \subseteq L, \mathcal{T}(\vee\{a_j : j \in J\}, b) = \vee\{\mathcal{T}(a_j, b) : j \in J\}$ (respectively, $\mathcal{T}(a, \wedge\{b_j : j \in J\}) = \wedge\{\mathcal{T}(a, b_j) : j \in J\}$).

For example, let $L = [0,1]$. Then for all $a, b \in L$,
(i) Gödel t-norm $\mathcal{T}_G(a,b) = \min\{a,b\}$; and
(ii) Łukasiewicz t-norm $\mathcal{T}_L(a,b) = \max\{0, a+b-1\}$.

Definition 3. *[2] A triangular conorm (t-conorm, for short) $\mathcal{S}$ is a map $\mathcal{S} : L \times L \to L$ such that the following conditions hold:*
(i) $\mathcal{S}(0,a) = a, \forall a \in L$;
(ii) $\mathcal{S}(a,b) = \mathcal{S}(b,a), \forall a, b \in L$;
(iii) $\mathcal{S}(a, \mathcal{S}(b,c)) = \mathcal{S}(\mathcal{S}(a,b), c), \forall a, b, c \in L$; and
(iv) if $a \le c$ and $b \le d$, then $\mathcal{S}(a,b) \le \mathcal{S}(c,d)$.
In addition, $\mathcal{S}$ is called a left (respectively, right) continuous if for all $a, b \in L$ and $\{a_j : j \in J\}, \{b_j : j \in J\} \subseteq L, \mathcal{S}(\wedge\{a_j : j \in J\}, b) = \wedge\{S(a_j, b) : j \in J\}$ (respectively, $\mathcal{S}(a, \vee\{b_j : j \in J\}) = \vee\{\mathcal{S}(a, b_j) : j \in J\}$.
For example, let $L = [0,1]$. Then for all $a, b \in L$,
(i) Gödel t-conorm $\mathcal{S}_G(a,b) = \max\{a,b\}$; and
(ii) Łukasiewicz t-conorm $\mathcal{S}_L(a,b) = \min\{1, a+b\}$.

Definition 4. *[2] A negator $\mathcal{N}$ is a decreasing map $\mathcal{N} : L \to L$ such that $\mathcal{N}(0) = 1$ and $\mathcal{N}(1) = 0$. If $\mathcal{N}(\mathcal{N}(a)) = a, \forall a \in L$, then $\mathcal{N}$ is called a strong negator.*
For a given negator $\mathcal{N}$, t-norm $\mathcal{T}$ and t-conorm $\mathcal{S}$ are called dual with respect to $\mathcal{N}$ if

$$\mathcal{S}(\mathcal{N}(a), \mathcal{N}(b)) = \mathcal{N}(\mathcal{T}(a,b)) \text{ and } \mathcal{T}(\mathcal{N}(a), \mathcal{N}(b)) = \mathcal{N}(\mathcal{S}(a,b)), \forall a, b \in L.$$

For example, let $L = [0,1]$. Then for all $a \in L$,
(i) standard negator $\mathcal{N}_{\mathcal{S}}(a) = 1 - a$, which is strong; and
(ii) Gödel negators

$$\mathcal{N}_{G_1}(a) = \begin{cases} 1 & if\ a = 0 \\ 0 & otherwise; \end{cases} \qquad \mathcal{N}_{G_2}(a) = \begin{cases} 0 & if\ a = 1 \\ 1 & otherwise, \end{cases}$$

which are non-strong least and greatest negators, respectively.

Definition 5. *[2] An implicator $\mathcal{I}$ is a map $\mathcal{I} : L \times L \to L$ such that $\mathcal{I}(1,0) = 0$ and $\mathcal{I}(0,0) = \mathcal{I}(0,1) = \mathcal{I}(1,1) = 1$. Further, $\mathcal{I}$ is called a left monotonic and right monotonic if for all $a \in L$, $\mathcal{I}(-,a)$ is a monotonically decreasing map and $\mathcal{I}(a,-)$ is a monotonically increasing map, respectively. Furthermore, $\mathcal{I}$ is called a hybrid monotonic if it is both left and right monotonic.*
In addition, $\mathcal{I}$ is called a left (respectively, right) continuous if for all $a, b \in L$ and $\{a_j : j \in J\}, \{b_j : j \in J\} \subseteq L, \mathcal{I}(\vee\{a_j : j \in J\}, b) = \wedge\{\mathcal{I}(a_j, b) : j \in J\}$ (respectively, $\mathcal{I}(a, \wedge\{b_j : j \in J\}) = \wedge\{\mathcal{I}(a, b_j) : j \in J\}$). For a left monotonic implicator $\mathcal{I}, \mathcal{N}_{\mathcal{I}}(a) = \mathcal{I}(a,0), \forall a \in L$ is called a negator induced by $\mathcal{I}$. Further, for $\lambda, \mu \in L^X, \lambda \rightsquigarrow \mu \in L$, is defined as:

$$\lambda \rightsquigarrow \mu = \wedge\{\mathcal{I}(\lambda(x), \mu(x)) : x \in X$$

For example, let $L = [0,1]$. Then for all $a, b \in L$,
(i) Łukasiewicz implicator $\mathcal{I}_L(a,b) = min\{1, 1-a+b\}$;
(ii) Kleene-Dienes implicator $\mathcal{I}_{KD}(a,b) = \max\{b, 1-a\}$;
(iii) Zadeh implicator $\mathcal{I}_{ZD}(a,b) = \max\{1-a, \min\{a,b\}$; and
(iv) Gödel implicator

$$\mathcal{I}_G(a,b) = \begin{cases} 1 & if\ a \le b \\ b & otherwise. \end{cases}$$

Definition 6. *[2] Let $\mathcal{I} : L \times L \to L$ be an implicator. Then $\mathcal{I}$ is called*
(i) a border implicator if $\mathcal{I}(1,a) = a, a \in L$;
(ii) a CP-implicator (CP stands for confinement principle) if $a \le b \Leftrightarrow \mathcal{I}(a,b) = 1, \forall a, b \in L$;
(iii) an EP-implicator (EP stands for exchange principle) if $\mathcal{I}(a, \mathcal{I}(b,c)) = \mathcal{I}(b, \mathcal{I}(a,c)), \forall a, b, c \in L$; and
(iv) an EN-implicator with t-norm $\mathcal{T}$ (EN stands for exchange principle with t-norm) if $\mathcal{I}(a, \mathcal{I}(b,c)) = \mathcal{I}(\mathcal{T}(a,b), c), \forall a, b, c \in L$.

Definition 7. *[1] Let $\tilde{F} = (Q, \Sigma, \tilde{R}, Z, \tilde{\delta}, w, F_1, F_2)$ be a general fuzzy automaton. If we fix an input $a_k \in \Sigma$ at time t_i the proposition $\alpha|_{a_k}$ can be computed by $\mu^{t_i}(q_i)$ if the general fuzzy automaton $\tilde{F}$ is in the state q_i at time t_i otherwise $\alpha|_{a_k}$ is 0 if $\tilde{F}$ is not in the active state q_i. Accordingly, for each state $q_i \in Q$ we can assess the truth value of $\alpha|_{a_k}$, it is indicated by $\alpha|_{a_k}(q_i)$. As explained above $\alpha|_{a_k}(q_i) \in [0,1]$. In this section, we derive the logic B which is a set of propositions about the general fuzzy automaton $\tilde{F}$ formulated by the observer and constructing a complete infinitely distributive lattice $\boldsymbol{B} = (B, \le, \wedge, \vee, 0, 1)$. We can establish the order $\le$ on B as follows:*
For $\alpha, \beta \in B, \alpha \le \beta$ if and only if $\alpha(q_i) \le \beta(q_i)$ for all $q_i \in Q$. One can instantly check that the contradiction, i.e., the proposition with constant truth value 0, is the least element and the tautology, i.e., the proposition with constant truth value 1, is the greatest component of the $\boldsymbol{B}$. Note that any component ith of 1 is the maximum membership values of active states at time t_i, for any $i \ge 0$.
We define L^B-valued subset of $Q \times \Sigma \times Q$, i.e., a map $\delta : Q \times \Sigma \times Q \to L^B$. The range set L^B allows to interpret L^B as a map assigning each (q, a_k, p) to $\delta_{a_k}(q,p) : B \to L$. This interpretation of transition map δ allows to represent it as the family $\{\delta^\alpha : \alpha \in B\}$ of L-valued sets $\delta^\alpha \in L^{Q \times \Sigma \times Q}$ of $Q \times \Sigma \times Q$ ordered by the elements of B, where the L-valued sets δ^α are defined by $\delta^\alpha_{a_k}(q,p) = \delta_{a_k}(q,p)(\alpha) = \begin{cases} 1 & if\ q = p \\ \alpha(q) \vee \alpha(p) & otherwise. \end{cases}$
An L^B-valued general fuzzy automaton is a 8-tuple $\tilde{F} = (Q, \Sigma, \tilde{R}, Z, \omega, \tilde{\delta}, F_1, F_2)$, where
(i) Q is a finite set of states, $Q = \{q_1, q_2, \ldots, q_n\}$,
(ii) Σ is a finite set of input symbols, $\Sigma = \{a_1, a_2, \ldots, a_m\}$,
(iii) $\tilde{R}$ is the set of fuzzy start states, $\tilde{R} \subseteq P(Q)$,
(iv) Z is a finite set of output symbols, $Z = \{b_1, b_2, \ldots, b_k\}$,
(v) $\omega : Q \to Z$ is the output function,
(vi) $\tilde{\delta} : (Q \times L) \times \Sigma \times Q \to L^B$ is the L^B-valued augmented transition function defined by

$$\tilde{\delta}^\alpha_{a_k}((q_i, \mu^t(q_i)), q_j) = F_1(\mu^t(q_i), \delta^\alpha_{a_k}(q_i, q_j)).$$

(vii) Function $F_1 : [0,1] \times [0,1] \to [0,1]$ is called membership assignment function.
(viii) Function $F_2 : [0,1]^* \times [0,1] \to [0,1]$ is called multi-membership resolution function.

Definition 8. *[1] Let $\tilde{F}$ be an L^B-valued general fuzzy automaton and $\alpha \in B$. Then the L^B-valued operators with*

t-norm are maps $\mathcal{T}R, \mathcal{T}R^{-1} : B \to (L^B)^Q$ *such that* $\forall q \in Q_{act}(t_i)$ *and* $\forall \lambda \in B$,

$$\mathcal{T}^{\alpha}R(\lambda)(q) = \vee\big\{\mathcal{T}(\lambda(p), \tilde{\delta^*}_u^{\alpha}((p, \mu^{t_j}(p)), q))| \\ p \in Q_{pred}(q,u), u \in \Sigma^*\big\};\ \textit{and}$$

$$\mathcal{T}^{\alpha}R^{-1}(\lambda)(q) = \vee\big\{\mathcal{T}(\lambda(p), \tilde{\delta^*}_u^{\alpha}((q, \mu^{t_i}(q)), p))| \\ p \in Q_{succ}(q,u), u \in \Sigma^*\big\}.$$

Definition 9. *[1] Let* $\tilde{F}$ *be an* L^B*-valued general fuzzy automaton. Then the* L^B*-valued operators with t-conorm are maps* $\mathcal{T}C, \mathcal{T}C^{-1} : B \to (L^B)^Q$ *such that* $\forall \lambda, \alpha \in B$ *and* $\forall q \in Q_{act}(t_i)$,

$$\mathcal{T}^{\alpha}C(\lambda)(q) = \wedge\big\{\mathcal{S}(\lambda(p), \mathcal{N}(\tilde{\delta^*}_u^{\alpha}((p, \mu^{t_j}(p)), q)))| \\ p \in Q_{pred}(q,u), u \in \Sigma^*\big\};$$

and

$$\mathcal{T}^{\alpha}C^{-1}(\lambda)(q) = \wedge\big\{\mathcal{S}(\lambda(p), \mathcal{N}(\tilde{\delta^*}_u^{\alpha}((q, \mu^{t_i}(q)), p)))| \\ p \in Q_{succ}(q,u), u \in \Sigma^*\big\}.$$

Definition 10. *[1] Let* $\tilde{F}$ *be an* L^B*-valued general fuzzy automaton and* $\alpha \in B$*. Then the* L^B*-valued operators with implicator are maps* $\mathcal{I}M, \mathcal{I}M^{-1} : B \to (L^B)^Q$ *such that* $\forall \lambda \in B$ *and* $\forall q \in Q_{act}(t_i)$,

$$\mathcal{I}^{\alpha}M(\lambda)(q) = \wedge\big\{\mathcal{I}(\tilde{\delta^*}_u^{\alpha}((p, \mu^{t_j}(p)), q), \lambda(p))| \\ p \in Q_{pred}(q,u), u \in \Sigma^*\big\};$$

and

$$\mathcal{I}^{\alpha}M^{-1}(\lambda)(q) = \wedge\big\{\mathcal{I}(\tilde{\delta^*}_u^{\alpha}((q, \mu^{t_i}(q)), p), \lambda(p))| \\ p \in Q_{succ}(q,u), u \in \Sigma^*\big\}.$$

Definition 11. *[2] A* ***Kuratowski*** *L****-fuzzy closure operator*** c *on a nonempty set* X *is a map* $c : L^X \to L^X$ *satisfying for all* $a \in L$, $\lambda, \mu \in L^X$, *and* $\{\lambda_j : j \in J\} \subseteq L^X$,
(i) $c(\boldsymbol{a}) = \boldsymbol{a}$;
(ii) $\lambda \leq c(\lambda)$;
(iii) $c(\lambda \vee \mu) = c(\lambda) \vee c(\mu)$; *and*
(iv) $c(c(\lambda)) = c(\lambda)$.
Further, Kuratowski L-fuzzy closure operator $c : L^X \to L^X$ *is called*
(v) Alexandroff if $c(\vee\{\lambda_j : j \in J\}) = \vee\{c(\lambda_j) : j \in J\}$.

Definition 12. *[2] A* ***Kuratowski*** *L****-fuzzy interior operator*** i *on a nonempty set* X *is a map* $i : L^X \to L^X$ *satisfying for all* $a \in L$, $\lambda, \mu \in L^X$, *and* $\{\lambda_j : j \in J\} \subseteq L^X$,
(i) $i(\boldsymbol{a}) = \boldsymbol{a}$;
(ii) $i(\lambda) \leq \lambda$;
(iii) $i(\lambda \wedge \mu) = i(\lambda) \wedge i(\mu)$; *and*
(iv) $i(i(\lambda)) = i(\lambda)$.
Further, Kuratowski L-fuzzy interior operator $i : L^X \to L^X$ *is called*
(v) ***Alexandroff*** *if* $i(\wedge\{\lambda_j : j \in J\}) = \wedge\{i(\lambda_j) : j \in J\}$.

Definition 13. *[2] An L-****fuzzy co-topology*** η *on a nonempty set* X *is a family of L-fuzzy sets in* X *which contains all constant L-fuzzy sets and is closed under arbitrary intersections and finite unions. The L-fuzzy sets in* η *are called closed, and the tuple* (X, η) *is called an L-fuzzy co-topological space. Further, the L-fuzzy co-topology* η *is called* ***Alexandroff*** *if an abitrary union of closed L-fuzzy sets is closed.*

Definition 14. *[2] An L-****fuzzy topology*** τ *on a nonempty set* X *is a family of L-fuzzy sets in* X *which contains all constant L-fuzzy sets and is closed under arbitrary unions and finite intersections. The L-fuzzy sets in* τ *are called open, and the tuple* (X, τ) *is called an L-fuzzy topological space. Further, the L-fuzzy topology* τ *is called* ***Alexandroff*** *if an arbitrary intersection of open L-fuzzy sets is open.*

III. L^B-VALUED TOPOLOGICAL CHARACTERIZATIONS OF L^B-VALUED GENERAL FUZZY AUTOMATA

Herein, we demonstrate that the L^B-valued operators with t-norm and the L^B-valued operators with t-conorm/implicator are Alexandroff Kuratowski L^B-valued closure operators and Alexandroff Kuratowski L^B-valued operators, respectively. Further, we demonstrate that the Alexandroff Kuratowski L^B-valued closure operators induce L^B-valued co-topologies and the Alexandroff Kuratowski L^B-valued interior operators induce L^B-valued topologies on state-set of a given L^B-valued general fuzzy automaton. Next, we demonstrate that the induced L^B-valued co-topologies/topologies can be used to characterize the algebraic concepts associated with an L^B-valued general fuzzy automaton. We begin with the following.

Proposition 1. *Let* $\tilde{F} = (Q, \Sigma, \tilde{R}, Z, \omega, \tilde{\delta}, F_1, F_2)$ *be an* L^B*-valued general fuzzy automaton an* $\mathcal{T}$ *is a Left continuous. Then*
(i) the L^B*-valued operators* $\mathcal{T}R, \mathcal{T}R^{-1} : B \to (L^B)^Q$ *are Alexandroff Kuratowski* L^B*-valued closure operators on* Q*; and*
(ii) these two L^B*-valued operators induce two* L^B*-valued co-topologies on* Q*, which we shall denote by* $\eta(Q)$ *and* $\eta^{-1}(Q)$*, respectively, and are given as follow:*

$$\eta(Q) = \{\lambda \in B : \mathcal{T}R(\lambda) = \lambda\};\ \textit{and}$$

$$\eta^{-1}(Q) = \{\lambda \in B : \mathcal{T}R^{-1}(\lambda) = \lambda\}.$$

Proof. Follows from Definition 8 and 11.

Proposition 2. *The* L^B*-valued co-topologies* $\eta(Q)$ *and* $\eta^{-1}(Q)$ *are Alexandroff.*

Proof. Let $\lambda_j \in \eta(Q), \forall j \in J$. Then $\mathcal{T}R(\lambda_j) = \lambda_j$. Now, $\mathcal{T}R(\lambda_j) = \lambda_j \Rightarrow \mathcal{T}R(\vee\{\lambda_j | j \in J\}) = \vee\{\mathcal{T}R(\lambda_j) | j \in J\} = \vee\{\lambda_j | j \in J\} \Rightarrow \vee\{\lambda_j | j \in J\} \in \eta(Q)$. Thus $\eta(Q)$ is an Alexandroff topology.
Similarly, we can show that the L^B-valued co-topology $\eta^{-1}(Q)$ is also an Alexandroff.

Proposition 3. *Let* $\tilde{F} = (Q, \Sigma, \tilde{R}, Z, \omega, \tilde{\delta}, F_1, F_2)$ *be an* L^B*-valued general fuzzy automaton,* $\mathcal{S}$ *be left continuous, and* $\mathcal{T}$ *and* $\mathcal{S}$ *be dual with respect to a strong negation* $\mathcal{N}$*. Then*
(i) The L^B*-valued operators* $\mathcal{T}C, \mathcal{T}C^{-1} : B \to (L^B)^Q$ *are Alexandroff Kuratowski* L^B*-valued interior operators on* Q*; and*

(ii) these two L^B-valued operators induce two L^B-valued topologies on Q, which we shall denote by $\tau(Q)$ and $\tau^{-1}(Q)$, respectively, and are given as follow:

$$\tau(Q) = \{\lambda \in B : \mathcal{T}C(\lambda) = \lambda\};\ \textit{and}$$

$$\tau^{-1}(Q) = \{\lambda \in B : \mathcal{T}C^{-1}(\lambda) = \lambda\}.$$

Proof. Follows from Definition 9 and 12.

Proposition 4. *The L^B-valued topologies $\tau(Q)$ and $\tau^{-1}(Q)$ are Alexandroff.*

Proof. Let $\lambda_j \in \tau(Q), \forall j \in J$. Then $\mathcal{T}C(\lambda_j) = \lambda_j$. Now, $\mathcal{T}C(\lambda_j) = \lambda_j \ \Rightarrow\ \mathcal{T}C(\wedge\{\lambda_j | j \in J\}) = \wedge\{\mathcal{T}C(\lambda_j) | j \in J\} = \wedge\{\lambda_j | j \in J\} \ \Rightarrow \wedge\{\lambda_j | j \in J\} \in \tau(Q)$. Thus $\tau(Q)$ is an Alexandroff topology.
Similarly, we can show that the L^B-valued topology $\tau^{-1}(Q)$ is also an Alexandroff.

Proposition 5. *Let $\eta(Q)$, $\eta^{-1}(Q)$ and $\tau(Q)$, $\tau^{-1}(Q)$ be L^B-valued co-topologies and L^B-valued topologies on Q, respectively. Then*
(i) $\eta(Q)$ and $\tau(Q)$ are dual in the sense that $\lambda \in B$ is an $\eta(Q)$-closed iff $\mathcal{N}(\lambda)$ is a $\tau(Q)$-open.
(ii) $\eta^{-1}(Q)$ and $\tau^{-1}(Q)$ are dual in the sense that $\lambda \in B$ is an $\eta^{-1}(Q)$-closed iff $\mathcal{N}(\lambda)$ is a $\tau^{-1}(Q)$-open, provided that $\mathcal{T}$ and $\mathcal{S}$ are dual with respect to a strong negation $\mathcal{N}$.

Proof. (i) Let $\lambda \in L^Q$ is an $\eta(Q)$-closed. Then $\mathcal{T}R(\lambda) = \lambda$. Now,

$$\begin{aligned}\mathcal{T}R(\lambda) = \lambda \ &\Leftrightarrow \mathcal{N}(\mathcal{T}R(\lambda)) = \mathcal{N}(\lambda)\\ &\Leftrightarrow \mathcal{T}C(\mathcal{N}(\lambda)) = \mathcal{N}(\lambda)\\ &\Leftrightarrow \mathcal{N}(\lambda) \in \tau(Q).\end{aligned}$$

Thus $\mathcal{N}(\lambda)$ is a $\tau(Q)$-open.
(ii) Similarly, we can show that $\eta^{-1}(Q)$ and $\tau^{-1}(Q)$ are dual in the sense that $\lambda \in L^Q$ is an $\eta^{-1}(Q)$-closed iff $\mathcal{N}(\lambda)$ is a $\tau^{-1}(Q)$-open.

Proposition 6. *Let $\tilde{F}$ be an L^B-valued general fuzzy automaton, $\mathcal{T}$ and $\mathcal{S}$ be dual with respect to a strong negator $\mathcal{N}$. Then $\lambda \in B$ is an L^B-valued subproposition of $\tilde{F}$ iff $\mathcal{N}(\lambda)$ is a $\tau(Q)$-open.*

Proof. Let λ is an L^B-valued subproposition of $\tilde{F}$. Then $\mathcal{T}R(\lambda) \leq \lambda$. Now,

$$\begin{aligned}\mathcal{T}R(\lambda) \leq \lambda \ &\Leftrightarrow \mathcal{N}(\lambda) \leq \mathcal{N}(\mathcal{T}R(\lambda))\\ &\Leftrightarrow \mathcal{N}(\lambda) \leq \mathcal{T}C(\mathcal{N}(\lambda))\\ &\Leftrightarrow \mathcal{N}(\lambda) \in \tau(Q).\end{aligned}$$

Thus $\mathcal{N}(\lambda)$ is a $\tau(Q)$-open.

Proposition 7. *Let $\tilde{F} = (Q, \Sigma, \tilde{R}, Z, \omega, \tilde{\delta}, F_1, F_2)$ be an L^B-valued general fuzzy automaton and $\mathcal{I}$ be a left-right continuous, border implicator, and EN-implicator with t-norm $\mathcal{T}$. Then*
(i) the L^B-valued operators $\mathcal{I}M, \mathcal{I}M^{-1} : B \to (L^B)^Q$ are Alexandroff Kuratowski L^B-valued interior operators on Q; and
(ii) these two L^B-valued operators induce two L^B-valued topologies on Q, which we shall denote by $\zeta(Q)$ and $\zeta^{-1}(Q)$, respectively, and are given as follows:

$$\zeta(Q) = \{\lambda \in B : \mathcal{I}M(\lambda) = \lambda\};\ \textit{and}$$

$$\zeta^{-1}(Q) = \{\lambda \in B : \mathcal{I}M^{-1}(\lambda) = \lambda\}.$$

Proof. Follows from Definition 10 and 12.

Definition 15. *Let $\boldsymbol{A} = (A, \leq, \wedge, \vee, 0, 1)$ be a complete infinitely distributive lattice of propositions about the general fuzzy automaton $\tilde{F} = (Q, \Sigma, \tilde{R}, Z, w, \tilde{\delta},\ F_1, F_2)$ and $\boldsymbol{B} = (B, \leq, \wedge, \vee, 0, 1)$ be a complete infinitely distributive lattice of propositions about the general fuzzy automaton $\tilde{F}' = (Q', \Sigma, \tilde{R}', Z, w', \tilde{\delta}',\ F_1, F_2)$. A homomorphism from an L^A-valued GFA $\tilde{F}$ to an L^B-valued GFA $\tilde{F}'$ is a pair (f, g) of maps, where $f : Q \to Q'$ and $g : A \to B$ are functions such that*

$$(i)\tilde{\delta}'((f(q), \mu^t(f(q))), u, f(p))(g(\alpha)) \geq \tilde{\delta}((q, \mu^t(q)), u, p)(\alpha),$$

$$(ii) w(q) = z \Leftrightarrow w(f(q)) = z,$$
$$(iii) f(q_0) = q'_0,$$
for all $\alpha \in A$ and $p, q \in Q$.
The pair (f, g) is called a strong homomorphism if for all $\alpha \in A$
$\tilde{\delta}'((f(q), \mu^t(f(q))), u, f(p))(g(\alpha)) = \vee\{\tilde{\delta}((q, \mu^t(q)), u, p')(\alpha) | p' \in Q, f(p) = f(p')\}$.

Example 1. *Consider the general fuzzy automata $\tilde{F} = (Q, \Sigma, \tilde{R}, Z, \omega, \tilde{\delta}, F_1, F_2)$ and $\tilde{F}' = (Q', \Sigma, \tilde{R}', Z, \omega', \tilde{\delta}', F_1, F_2)$ in Figure 1 and Figure 2, respectively.*

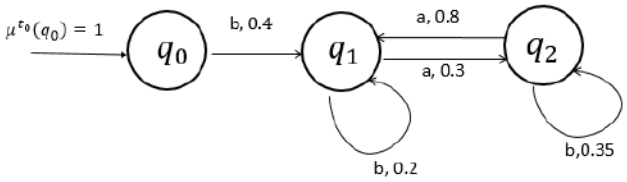


***Fig. 1** The GFA $\tilde{F}$ of Example 1*

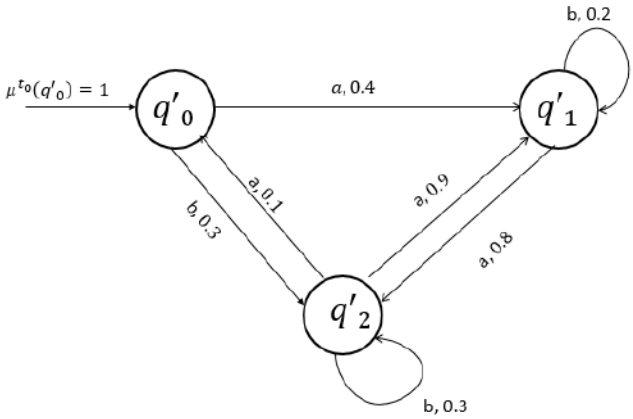


***Fig. 2** The GFA $\tilde{F}'$ of Example 1*

If we choose $F_1(\mu,\delta)=\delta, F_2()=\mu^{t+1}(q_m)=\wedge_{i=1}^{n}(F_1(\mu^t(q_i),\ \delta(q_i,a_k,q_m))$*, then the operations of these automata upon input string "*ba^2b*" are shown in Table 1 and 2, respectively. We have the set* $B=\{0,\alpha_0,\alpha_1,\alpha_2,\alpha_3,\alpha_4,1\}$

TABLE I
ACTIVE STATES AND THEIR MEMBERSHIP VALUES (MV) AT DIFFERENT TIMES OF $\tilde{F}$ IN EXAMPLE 1 UPON INPUT STRING "ba^2b"

time	t_0	t_1	t_2	t_3	t_4
input	$\wedge$	b	a	a	b
$Q_{act}(t_i)$	q_0	q_1	q_2	q_1	q_1
mv	1	0.4	0.3	0.8	0.2

TABLE II
ACTIVE STATES AND THEIR MEMBERSHIP VALUES (MV) AT DIFFERENT TIMES OF $\tilde{F}'$ IN EXAMPLE 1 UPON INPUT STRING "ba^2b"

time	t_0	t_1	t_2	t_3	t_4
input	$\wedge$	b	a	a	b
$Q_{act}(t_i)$	q'_0	q'_2	q'_0 | q'_1	q'_1 | q'_2	q'_1 | q'_2
mv	1	0.3	0.1 | 0.9	0.4 | 0.8	0.2 | 0.3

of possible propositions ***B*** *about the general fuzzy automaton* $\tilde{F}$ *and* $A=\{0',\beta_0,\beta_1,\beta_2,\beta_3,\beta_4,1'\}$ *of possible propositions* ***A*** *about the general fuzzy automaton* $\tilde{F}'$*, where* $0=(0,0,0),\alpha_0=(1,0,0),\alpha_1=(0,0.4,0),\alpha_2=(0,0,0.3),\alpha_3=(0,0.8,0),\alpha_4=(0,0.2,0),1=(1,0.8,0.3)$ *and* $0'=(0,0,0),\beta_0=(1,0,0),\beta_1=(0,0,0.3),\beta_2=(0.1,0.9,0),\beta_3=(0,0.4,0.8),\ \beta_4=(0,0.2,0.3),\ 1'=(1,0.9,0.8)$. *According to the definition of homomorphism between* L^B*-valued general fuzzy automata, we define a homomorphism* $f:\tilde{F}\to\tilde{F}'$ *with a pair* (f,g) *of maps, where* $f:Q\to Q'$ *defined by,* $f(q_i)=q'_i$ *and* $g:B\to A$ *defined by:* $g(0)=0',\ g(1)=1',\ g(\alpha_1)=\beta_3,\ g(\alpha_2)=\beta_1,\ g(\alpha_3)=\beta_2,\ g(\alpha_4)=\beta_4$.
For example we have:

$$\begin{aligned}\tilde{\delta}((q_1,\mu^{t_1}(q_1)),a,q_2)(\alpha_2)&=F_1(\mu^{t_1}(q_1),\delta(q_1,a,q_2)(\alpha_2))\\&=\delta(q_1,a,q_2)(\alpha_2)\\&=\alpha_2(q_1)\vee\alpha_2(q_2)=0\vee0.3=0.3,\end{aligned}$$

and $\tilde{\delta}'((f(q_1),\mu^{t_2}(f(q_1))),a,f(q_2))(g(\alpha_2))=\tilde{\delta}'((q'_1,\mu^{t_2}(q'_1)),a,q'_2)(\beta_1)=F_1(\mu^{t_2}(q'_1),\delta'(q'_1,a,q'_2)(\beta_1))=\delta'(q'_1,a,q'_2)(\beta_1)=\beta_1(q'_1)\vee\beta_1(q'_2)=0\vee0.3=0.3$*, then* $\tilde{\delta}'((f(q_1),\mu^{t_2}(f(q_1))),a,f(q_2))(g(\alpha_2))\geq\tilde{\delta}((q_1,\mu^{t_1}(q_1)),a,q_2)(\alpha_2)$*. Also,* $\tilde{\delta}((q_2,\mu^{t_2}(q_2)),a,q_1)(\alpha_1)=F_1(\mu^{t_2}(q_2),\delta(q_2,a,q_1)(\alpha_1))=\delta(q_2,a,q_1)(\alpha_1)=\alpha_1(q_2)\vee\alpha_1(q_1)=0\vee0.4=0.4$*, and* $\tilde{\delta}'((f(q_2),\mu^{t_1}(f(q_2))),a,f(q_1))(g(\alpha_1))=\tilde{\delta}'((q'_2,\mu^{t_1}(q'_2)),a,q'_1)(\beta_3)=F_1(\mu^{t_1}(q'_2),\delta'(q'_2,a,q'_1)(\beta_3))=\delta'(q'_2,a,q'_1)(\beta_3)=\beta_3(q'_2)\vee\beta_3(q'_1)=0.8\vee0.4=0.8$*, then* $\tilde{\delta}'((f(q_2),\mu^{t_1}(f(q_2))),a,f(q_1))(g(\alpha_1))\geq\tilde{\delta}((q_2,\mu^{t_2}(q_2)),a,q_1)(\alpha_1)$.

Definition 16. *Let* (X,τ) *and* (X',τ') *be the* L^B*-valued cotopological/topological spaces. A map* $\varphi:(X,\tau)\to(X',\tau')$ *is called an* L^B*-valued continuous if for each* $\lambda'\in\tau'$*,* $\varphi^{-1}(\lambda')\in\tau$.

Proposition 8. *If* $\varphi:\tilde{F}\to\tilde{F}'$ *is a homomorphism between* L^B*-valued general fuzzy automata, then* $\varphi:(Q,\eta(Q))\to(Q',\eta(Q'))$*,* $\varphi:(Q,\eta^{-1}(Q))\to(Q',\eta^{-1}(Q'))$*,* $\varphi:(Q,\tau(Q))\to(Q',\tau(Q'))$*,* $\varphi:(Q,\tau^{-1}(Q))\to(Q',\tau^{-1}(Q'))$*,* $\varphi:(Q,\zeta(Q))\to(Q',\zeta(Q'))$*,* $\varphi:(Q,\zeta^{-1}(Q))\to(Q',\zeta^{-1}(Q'))$ *are* L^B*-valued continuous.*

Proof. We only prove that $\varphi:(Q,\eta(Q))\to(Q',\eta(Q'))$ is an L^B-valued continuous, the proofs of the rest will follow the same process. To demonstrate that φ is an L^B-valued continuous, it is enough to show that $\lambda'\in\eta(Q')\Rightarrow\varphi^{\leftarrow}(\lambda')\in\eta(Q)$. For which, let $\lambda'\in\eta(Q')$. Then $\mathcal{T}^\alpha R(\lambda')=\lambda'$. Now, for all $q\in Q$.

$$\begin{aligned}&\mathcal{T}^\alpha R(\varphi^{\leftarrow}(\lambda'))(q)=\vee\{\mathcal{T}(\varphi^{\leftarrow}(\lambda')(p),\tilde{\delta^*}^{\alpha}_{u}((p,\mu^{t_j}(p)),q))|\\&\quad p\in Q_{pred}(q,u),u\in\Sigma^*\}\\&=\vee\{\mathcal{T}(\lambda'(\varphi(p)),\tilde{\delta^*}^{\alpha}_{u}((p,\mu^{t_j}(p)),q))\\&\quad |p\in Q_{pred}(q,u),u\in\Sigma^*\}\\&\leq\vee\{\mathcal{T}(\lambda'(\varphi(p)),\tilde{\delta^{*'}}^{\alpha}_{u}((\varphi(p),\mu^{t_j}(\varphi(p))),\varphi(q))\\&\quad |p\in Q_{pred}(q,u),u\in\Sigma^*\}\\&\leq\vee\{\mathcal{T}(\lambda'(p'),\tilde{\delta^{*'}}^{\alpha}_{u}((p',\mu^{t_j}(p')),\varphi(q)))\\&\quad |p'\in Q_{pred}(\varphi(q),u),u\in\Sigma^*\}\\&=\mathcal{T}^\alpha R(\lambda')(\varphi(q))\\&=\lambda'(\varphi(q))\\&=\varphi^{\leftarrow}(\lambda')(q)\end{aligned}$$

Thus $\mathcal{T}^\alpha R(\varphi^{\leftarrow}(\lambda'))\leq\varphi^{\leftarrow}(\lambda')$, which implies that $\varphi^{\leftarrow}(\lambda')\in\eta(Q)$. Hence φ is an L^B-valued continuous.

IV. CONCLUSION

This paper has demonstrated that the L^B-valued operators with t-norm and the L^B-valued operators with t-conorm/implicator are Alexandroff Kuratowski L^B-valued closure operators and Alexandroff Kuratowski L^B-valued interior operators, respectively. Interestingly, the L^B-valued cotopologies and L^B-valued topologies on state-set of a given L^B-valued general fuzzy automaton are introduced and shown to be useful for the characterization of algebraic concepts of such automaton.

REFERENCES

[1] K. Abolpour, On L^B-valued GFA: An L^B-valued operator oriented view with t-norm/t-conorm and implicators, New Mathematics and Natural Computation, In Press.

[2] D. Qiu, C. Lu, W. Zhang and Y. Lan, Algebraic properties and topological properties of the quotient space of fuzzy numbers based on Mares equivalence relation, Fuzzy Sets and Systems, 245 (2014) 63-82.

[3] E.S. Santos, Maximin automata, Information and Control, 12 (1968) 367-377.

[4] S.P. Tiwari and A.K. Srivastava, A topology for fuzzy automata, AFSS International Conference on Fuzzy Systems, 2275 (2002) 485-491.

[5] W.G. Wee, On generalizations of adaptive algorithm and application of the fuzzy sets concept to pattern classification, Ph. D. Thesis, Purdue University, (1967).

[6] W.G. Wee and K.S. Fu, A formulation of fuzzy automata and its application as a model of learning systems,

IEEE Transaction on Systems Science and Cybernetics, 5 (1969) 215-223.
[7] L.A. Zadeh, Fuzzy sets, Information and Control, 8 (1965) 338-353.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

The Decomposition Theorems for Residuated Lattices via Directed Kernels

1st Farideh Frasad, 2nd Mohammad Ali Nourollahi

1st Department of Pure Mathematics, Faculty of Mathematics and Computer, Higher Education Complex of Bam, Bam, Iran, faridehfarsad@yahoo.com
2nd Department of Pure Mathematics, Faculty of Mathematics and Computer, Higher Education Complex of Bam, Bam, Iran, mnourollahi@bam.ac.ir

***Abstract*— The isomorphism theorems are those that describe the relationship between quotients, homomorphisms, and subobjects. Also, to construct the quotient of residuated lattices, whose algebraic structures is determined by a partial order, it is more useful to consider directed kernels of homomorphisms between such algebras. So, in this paper, we use the concept of directed kernel of a resituated lattice homomorphism and that of a pre-congruence. Finally, we prove the decomposition and isomorphism theorems for residuated lattice homomorphisms.**

***Index Terms*— Pre-congruence, Residuated lattice congruence, Residuated lattice, Decomposition Theorems.**

I. Introduction and Preliminaries

The interest in lattice-valued logic has been rapidly growing recently. Several algebras playing the role of structures of true values have been introduced and axiomatized [3]. The most general structure considered in this paper is that of a residuated lattice.

Residuated lattices were first introduced as a generalization of ideal lattices of rings in 1939 by Ward and Dilworth[6].

The isomorphism theorems can be generalized to the context of algebras and congruences. In [4], the authors take a close look at in directed kernels of residuated lattice homomorphisms and characterize them as pre-congruences on a residuated lattices. So in this paper by these notions, we prove the counterpart of algebra's first, second and third isomorphism theorems for residuated lattices.

Let's now recall, for example from [6], the notion of residuated lattices and residuated lattice homomorphisms.

Definition 1. *A residuated lattice is an algebra* $(P, \wedge, \vee, \odot, \rightarrow, 0, 1)$ *of type (2,2,2,2, 0,0) equipped with an order* $\leq$ *satisfying the following:*
(LR_1) $(P, \wedge, \vee, 0, 1)$ *is a bounded lattice,*
(LR_2) $(P, \odot, 1)$ *is a commutative ordered monoid,*
(LR_3) $\odot$ *and* $\rightarrow$ *form an adjoint pair, i.e.,* $c \leq a \rightarrow b$ *iff* $a \odot c \leq b$*, for all* $a, b, c \in P$*.*

The relations between the pair of operations $\odot$ and $\rightarrow$ expressed by LR_3, is a particular case of the law of residuation, or Galois correspondence (see [3]) and for every $x, y \in A$, $x \rightarrow y = \sup\{z \in A \mid x \odot z \leq y\}$.

A function $f : P \rightarrow Q$ between two residuated lattices P and Q is a *residuated lattice homomorphism* if it is a morphism of bounded lattices and for every $x, y \in P$: $f(x \odot y) = f(x) \odot f(y)$ and $f(x \rightarrow y) = f(x) \rightarrow f(y)$.

Example 1. *If on* $I = [0, 1]$*; for* $x, y \in I$ *we define* $x \odot y = \min\{x, y\}$ *and* $x \rightarrow y = 1$ *if* $x \leq y$ *and* y *otherwise, then* $(I, \max, \min, \odot, \rightarrow 0, 1)$ *is a residuated. lattice.*

II. Pre-congruence relations on Residuated lattices

First of all, we recall some concepts of pre-ordered sets which will be used in the sequel (see [3]).
Consider a set P with a reflexive and transitive relation σ. Such a relation will be called a *pre-order* and (P, σ) *a pre-ordered set*.

Let (P, σ) be a pre-ordered set. Then the relation $\overline{\sigma}$ defined by $p \ \overline{\sigma} \ q$ if and only if $p \ \sigma \ q$ and $q \ \sigma \ p$, for each $p, q \in P$, is an equivalence relation on P. Moreover, it is the largest equivalence relation on P with the property that $\overline{\sigma} \subseteq \sigma$.

It is a well-known fact that $(P/\overline{\sigma}, \sigma/\overline{\sigma})$ is a poset. Moreover, $\overline{\sigma}$ is the smallest equivalence relation with this property. We also have that for every $X, Y \in P/\overline{\sigma}$, $X \ (\sigma/\overline{\sigma}) \ Y$ if and only if $x \ \sigma \ y$ for some $x \in X$ and $y \in Y$ if and only if $x \ \sigma \ y$ for each $x \in X$ and $y \in Y$.

Definition 2. *Let* $(P, \leq)$ *be a pre-ordered set and let* $A \subseteq P$*. An element* $x \in P$ *is a pre-upper bound of* A *if* $a \leq x$*, for all* $a \in A$*. A pre-lower bound is defined dually. In other words, an element* $y \in P$ *is a pre-lower bound of* A *if* $y \leq a$*, for all* $a \in A$*.*

Definition 3. *Let* $(P, \leq)$ *be a pre-ordered set and let* $A \subseteq P$*. An element* $x \in P$ *is called a pre-supremum of* A *if*

(i) *x is a pre-upper bound of A, and*
(ii) *$x \leq y$ for all pre-upper bounds y of A.*

A pre-infimum of A is also defined dually.

In the following, we recall from [4], a pre-congruence relation on a residuated lattice which has an important role in characterizing residuated lattice congruences.

Definition 4. *Let P and Q be two pre-complete pre-ordered sets and $f\colon P \to Q$ be a map.*

(i) *f preserves arbitrary pre-suprema if s is a pre-supremum of $A \subseteq P$ then $f(s)$ is a pre-supremum of $f(A)$ in Q, for all $A \subseteq P$.*
(ii) *f preserves arbitrary pre-infima if s is a pre-infimum of $A \subseteq P$ then $f(s)$ is a pre-infimum of $f(A)$ in Q, for all $A \subseteq P$.*

Definition 5. *Let $(P, \leq)$ be a residuated lattice, a pre-order σ on P is called pre-congruence if it satisfies the following conditions:*

(1) $\leq \subseteq \sigma$.
(2) *(P, σ) has binary pre-suprema and binary pre-infima and $(P, \sigma, \odot)$ is a pomonoid.*
(3) *$\odot$ and $\to$ form an adjoint pair, i.e., $c\ \sigma\ (a \to b)$ iff $(a \odot c)\ \sigma\ b$, for all $a, b, c \in A$, where $a \to b =$ pre-suprema$\{c \in P \mid (a \odot c)\ \sigma\ b\}$.*
(4) *The identity function $id_P\colon (P, \leq) \longrightarrow (P, \sigma)$ preserves binary pre-suprema and binary pre-infima.*

Theorem 1. *Let $(P, \leq)$ be a residuated lattice and σ a pre-order on P. Then σ is a pre-congruence on P if and only if*

(i) $\leq \subseteq \sigma$;
(ii) *If $a\ \sigma\ b$ and $a'\ \sigma\ b$ then $(a \vee a')\ \sigma\ b$, for all a, b, a' in A;*
(iii) *If $a\ \sigma\ b$ and $a'\ \sigma\ b$ then $(a \wedge a')\ \sigma\ b$, for all a, b, a' in A;*
(iv) *If $a\ \sigma\ b$ and $a'\ \sigma\ b$ then $(a \odot a')\ \sigma\ b$, for all a, b, a' in A;*
(v) *$\odot$ and $\to$ form an adjoint pair, i.e., $c\ \sigma\ (a \to b)$ iff $(a \odot c)\ \sigma\ b$, for all $a, b, c \in A$, where $a \to b =$ pre-suprema$\{c \in P \mid (a \odot c)\ \sigma\ b\}$.*

Definition 6. *Let $(P, \leq)$ and $(Q, \sqsubseteq)$ be posets and $f\colon P \to Q$ be an order-preserving map. The set*

$$\vec{ker} f := \{(a, b) \in P \times P \mid f(a) \sqsubseteq f(b)\},$$

is called the directed kernel or *sub-kernel of f.*

Theorem 2. *Let P and Q be residuated lattices, and $f\colon P \longrightarrow Q$ be a residuated lattice homomorphism. Then $\vec{\ker} f$ is a pre-congruence.*

The following theorem gives a characterization of pre-congruences as the directed kernels of residuated lattice homomorphisms.

Theorem 3. *Let $(P, \leq)$ be a residuated lattice and σ be a pre-order on P. Then the following are equivalent:*

1. *σ is a pre-congruence on P.*
2. *$(P/\overline{\sigma}, \sigma/\overline{\sigma})$ is a residuated lattice, the canonical surjection $\pi\colon (P, \leq) \longrightarrow (P/\overline{\sigma}, \sigma/\overline{\sigma})$ is a residuated lattice homomorphism and $\vec{\ker}\pi = \sigma$.*
3. *There exist a residuated lattice $(Q, \preceq)$ and a residuated lattice homomorphism $f\colon (P, \leq) \longrightarrow (Q, \preceq)$ such that $\vec{\ker} f = \sigma$.*

The final theorem characterizes residuated lattice congruences via a pre-congruence.

Definition 7. *An equivalence relation R on a residuated lattice $(P, \leq)$ is said to be a residuated lattice congruence if there is an order $\preceq$ on P/R such that $(P/R, \preceq)$ is a residuated lattice and the canonical map $\pi_R\colon P \to P/R$, $p \mapsto [p]_R$, is a residuated lattice map.*

Theorem 4. *Let $(P, \leq)$ be a residuated lattice and R be an equivalence relation on P. Then the following are equivalent:*

1. *R is a residuated lattice congruence.*
2. *There exists a pre-congruence σ on P such that $R = \sigma \cap \sigma^{-1}$.*
3. *There exist a residuated lattice Q and a residuated lattice homomorphism $f\colon P \to Q$ such that $R = \ker f$.*

III. The Decomposition Theorem for Residuated lattices

In this section, first of all, we state and prove the Decomposition Theorem and as its consequences for residuated lattices. We also prove the counterpart of algebra's first, second and third isomorphism theorems for residuated lattices.

Theorem 5 (The Decomposition Theorem). *Let $f\colon P \to Q$ and $g\colon P \to T$ be residuated lattice homomorphisms such that*

(1) *f is surjective,*
(2) $\vec{\ker} f \subseteq \vec{\ker} g$.

Then there is a unique residuated lattice homomorphism $h\colon Q \to T$ such that

$$h \circ f = g.$$

Moreover, h is injective if and only if $\vec{\ker} f = \vec{\ker} g$. Also, h is surjective if and only if g is surjective.

Proof: By Theorem 2, $\vec{\ker} f$ and $\vec{\ker} g$ are pre-congruences. Theorem 3 implies that $\ker f$ and $\ker g$ are residuated lattice congruences. Also, by (2), $\vec{\ker} f \subseteq \vec{\ker} g$, and so

$$\ker f = (\vec{\ker} f) \cap (\vec{\ker} f)^{-1} \subseteq (\vec{\ker} g) \cap (\vec{\ker} g)^{-1} = \ker g.$$

Now, since f is surjective and $\ker f \subseteq \ker g$, by the Decomposition Theorem of functions, the map $h\colon Q \to T$ defined by $h(q) = g(p_q)$, where $p_q \in P$ is an element such that $f(p_q) = q$, is the unique function satisfying $h \circ f = g$. Moreover, h is surjective if and only if g is surjective.

Now, we show that h is a residuated lattice homomorphism. First notice that h preserves finite joins. To see this, let A be a finite subset of Q. Then, we show that $h(\bigvee A) = \bigvee h(A)$. Take $r := \bigvee A \in Q$ and $s := \bigvee f^{-1}(A) \in P$. So $f(s) = f(\bigvee f^{-1}(A)) = \bigvee f(f^{-1}(A)) = \bigvee A$ (notice that since f is surjective, $f(f^{-1}(A)) = A$). This gives that $h(\bigvee A) = g(s)$. We also have $h(A) = g(f^{-1}(A))$ and so $g(s) = g(\bigvee f^{-1}(A)) = \bigvee g(f^{-1}(A)) = \bigvee h(A)$.

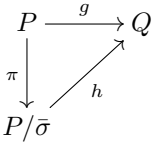
Consequently, $h(\bigvee A) = g(s) = \bigvee h(A)$, as required.
Now, we show that h preserves finite meets. To prove this, let B be a finite subset of Q. Take $z := \bigwedge B \in Q$ and $t := \bigvee f^{-1}(B) \in P$. So $f(t) = f(\bigwedge f^{-1}(B)) = \bigwedge f(f^{-1}(B)) = \bigwedge B$ (notice that $f(f^{-1}(B)) = B$, since f is surjective). Thus, $h(\bigwedge B) = g(t)$. We also have $h(B) = g(f^{-1}(B))$ and so $g(t) = g(\bigwedge f^{-1}(B)) = \bigwedge g(f^{-1}(B)) = \bigwedge h(B)$. Consequently, $h(\bigwedge B) = g(t) = \bigwedge h(B)$, as required.
Also, h preserves two operations $\odot$ and $\to$. In fact, since f is surjective so for $a, b \in Q$, we have $h(a) = hf(f^{-1}(a)) = g(f^{-1}(a))$, then $h(a \odot b) = hf(f^{-1}(a) \odot f^{-1}(b)) = g(f^{-1}(a) \odot f^{-1}(b)) = gf^{-1}(a) \odot gf^{-1}(b) = h(a) \odot h(b)$. By similar proof, we get $h(a \to b) = h(a) \to h(b)$.

Finally, suppose that $\vec{\ker} f = \vec{\ker} g$. We show that h is injective. To see this, let $h(q_1) = h(q_2)$ for $q_1, q_2 \in Q$. Since f is surjective, there exist $p_1, p_2 \in P$ such that $f(p_1) = q_1$, $f(p_2) = q_2$. Thus $g(p_1) = hf(p_1) = h(q_1) = h(q_2) = hf(p_2) = g(p_2)$. So, by hypothesis, $f(p_1) = f(p_2)$, that is $q_1 = q_2$.

Conversely, if $h\colon Q \to g(P)$ is injective, then $\vec{\ker} g = \vec{\ker} f$. By (2), $\vec{\ker} f \subseteq \vec{\ker} g$. To see $\vec{\ker} g \subseteq \vec{\ker} f$, let $x \ \vec{\ker} g \ y$. Then $g(x) \leq_T g(y)$ and so $g(x \wedge y) = g(x) \wedge g(y) = g(y)$. Thus $h(f(x \wedge y)) = g(x \wedge y) = g(y) = h(f(y))$. Therefore, $f(x) \wedge f(y) = f(x \wedge y) = f(y)$, since h is injective. So $f(x) \leq f(y)$. This gives that $x \ \vec{\ker} f \ y$, as required. ∎

As a consequence, we have:

Corollary 1. *Let $g\colon P \to Q$ be a residuated lattice homomorphism. Also, let σ be a pre-congruence on P that $\sigma \subseteq \vec{\ker} g$. Then there exists a unique residuated lattice homomorphism $h\colon P/\bar{\sigma} \to Q$ which makes the diagram*



commutative. Moreover, h is injective if and only if $\sigma = \vec{\ker} g$. Also, h is surjective if and only if g is surjective.

Applying Theorem 4, we obtain:

Corollary 2 (The First Isomorphism Theorem). *Let $f\colon P \to Q$ be a residuated lattice homomorphism. Then*

$$P/\ker f \cong f(P).$$

In particular, if $f\colon P \to Q$ is a surjective residuated lattice homomorphism, then

$$P/\ker f \cong Q.$$

To prove the second isomorphism theorem, we first give a description of pre-congruences on the quotient residuated lattices.

Remark 1. *(1) If $f\colon (P, \leq) \to (Q, \sqsubseteq)$ is a surjective residuated lattice homomorphism and σ is a pre-congruence on P such that $\vec{\ker} f \subseteq \sigma$, then $(f \times f)(\sigma)$ is a pre-congruence on Q.*

(2) Let $(P, \leq)$ be a residuated lattice and σ, ρ two pre-congruences on P with $\rho \subseteq \sigma$. Then the quotient relation

$$\sigma/\rho := \{([a]_{\bar{\rho}}, [b]_{\bar{\rho}}) \mid (a, b) \in \sigma\}$$

is a pre-congruence on the residuated lattice $(P/\bar{\rho}, \rho/\bar{\rho})$. This is because, by Theorem 3, the canonical surjection $\pi_{\bar{\rho}}\colon (P, \leq) \to (P/\bar{\rho}, \rho/\bar{\rho})$ is a residuated lattice homomorphism, and $\vec{\ker}\pi_{\bar{\rho}} = \rho \subseteq \sigma$. So, by (1), $\sigma/\bar{\rho} = (\pi_{\bar{\rho}} \times \pi_{\bar{\rho}})(\sigma)$ is a pre-congruence on $(P/\bar{\rho}, \rho/\bar{\rho})$.

(3) Every residuated lattice congruence of a quotient residuated lattice P/R is of the form $\overline{\theta/R}$, where $R = \bar{\sigma}$ (see Theorem 4), for some pre-congruence σ on P, and θ is a pre-congruence on P containing σ.

This is because by (2), for each pair of pre-congruences θ and σ with $\sigma \subseteq \theta$, $\theta/\bar{\sigma}$ is a pre-congruence on $P/\bar{\sigma}$. So, by Theorem 4, $\overline{\theta/\bar{\sigma}}$ is a residuated lattice congruences on P/R. On the other hand, for each residuated lattice congruence Θ on the quotient residuated lattice $(P/\bar{\sigma}, \sigma/\bar{\sigma})$, by Theorem 4, $\Theta = \bar{\rho}$, for some pre-congruence ρ on the residuated lattice $(P/\bar{\sigma}, \sigma/\bar{\sigma})$. Now, $\theta := \{(x, y) \in P \times P \mid ([x]_{\bar{\sigma}}, [y]_{\bar{\sigma}}) \in \rho\}$ is a pre-congruence on P containing σ, also $\rho = \theta/\bar{\sigma}$, and so $\Theta = \overline{\theta/\bar{\sigma}}$.

To prove the second isomorphism theorem for residuated lattices, we need the following lemma.

Lemma 1. *Let $(P, \leq)$ be a residuated lattice and σ, ρ be two pre-congruences on P with $\rho \subseteq \sigma$. Then Pre-sup$_{(P,\rho)}(A) \subseteq$ Pre-sup$_{(P,\sigma)}(A)$ and also Pre-inf$_{(P,\rho)}(A) \subseteq$ Pre-inf$_{(P,\sigma)}(A)$, for all $A \subseteq P$.*

Proof: We have $\rho \subseteq \sigma$. This fact and $\bigvee A \in$ Pre-sup$_{(P,\rho)}(A) \cap$ Pre-sup$_{(P,\sigma)}(A)$ (see Theorem 3) imply that each pre-supremum of every subset A of (P, ρ) is also a pre-supremum of A in (Q, σ), as required. In the same way, one can easily show this fact for the pre-infima. ∎

Corollary 3 (The Second Isomorphism Theorem). *Let P be a residuated lattice, ρ, σ pre-congruences on P with $\rho \subseteq \sigma$. Then*

$$P/\bar{\rho}/\overline{\sigma/\bar{\rho}} \cong P/\bar{\sigma}.$$

Proof: First, we show that $f\colon (P/\bar{\rho}, \rho/\bar{\rho}) \to (P/\bar{\sigma}, \sigma/\bar{\sigma})$, $[p]_{\bar{\rho}} \mapsto [p]_{\bar{\sigma}}$, is a residuated lattice homomorphism. First, we show that f preserves finite joins. To see this, let $A \subseteq (P/\bar{\rho}, \rho/\bar{\rho})$ and $[t]_{\bar{\rho}} := \bigvee A$ where t is a pre-supremum of $\pi_{\bar{\rho}}^{-1}(A)$ in (P, ρ) (see the proof (1) ⇒ (2) in Theorem 3 from [4]). By Lemma 1, t is also a pre-supremum of $\pi_{\bar{\rho}}^{-1}(A)$ in (P, σ). Hence $[t]_{\bar{\sigma}} = f([t]_{\bar{\rho}})$ is a supremum of

$$\begin{aligned} \{[x]_{\bar{\sigma}} \mid x \in \pi_{\bar{\rho}}^{-1}(A)\} &= \{[x]_{\bar{\sigma}} \mid [x]_{\bar{\rho}} \in A\} \\ &= \{f([x]_{\bar{\rho}}) \mid [x]_{\bar{\rho}} \in A\} \\ &= f(A). \end{aligned}$$

Hence f preserves finite joins. In the same way, one can show that f preserves finite meets. Also, it is easy to check that f preserves $\to$ and $\odot$ and conclude that f is a residuated lattice homomorphism.

Finally, by the definition of $\vec{\ker} f$ and the quotient relations $\sigma/\bar{\sigma}$ and $\sigma/\bar{\rho}$, we have

$$\begin{aligned}([a]_{\bar{\rho}},[b]_{\bar{\rho}}) \in \vec{\ker} f &\Leftrightarrow [a]_{\bar{\sigma}}\ \sigma/\bar{\sigma}\ [b]_{\bar{\sigma}}\\ &\Leftrightarrow (a,b)\in\sigma\\ &\Leftrightarrow ([a]_{\bar{\rho}},[b]_{\bar{\rho}})\in\sigma/\bar{\rho}.\end{aligned}$$

So, $\ker f = \vec{\ker} f \cap (\vec{\ker} f)^{-1} = \sigma/\bar{\rho}\cap(\sigma/\bar{\rho})^{-1} = \overline{\sigma/\bar{\rho}}$. Now, by Corollary 2, we have $P/\bar{\rho}/\overline{\sigma/\bar{\rho}} \cong P/\bar{\sigma}$. ∎

To prove the third isomorphism theorem, we study the relation between the pre-congruences on a residuated lattice and its subresiduated lattices.

Lemma 2. *Let $(P,\leq)$ be a residuated lattice, Q a subresiduated lattice of P and σ a pre-congruence on $(P,\leq)$. Then $\sigma\restriction_Q = \sigma\cap(Q\times Q)$ is a pre-congruence on Q and $\overline{\sigma\restriction_Q} = \bar{\sigma}\cap(Q\times Q)$.*

Proof: By Theorem 1, since σ is a pre-congruence on P and Q is closed under binary joins and binary meets, $\sigma\restriction_Q$ is trivially a pre-congruence on Q. Moreover,

$$\begin{aligned}\overline{\sigma\restriction_Q} &= \sigma\restriction_Q\cap(\sigma\restriction_Q)^{-1}\\ &= (\sigma\cap(Q\times Q))\cap(\sigma\cap(Q\times Q))^{-1}\\ &= (\sigma\cap(Q\times Q))\cap(\sigma^{-1}\cap(Q\times Q))\\ &= (\sigma\cap\sigma^{-1})\cap(Q\times Q)\\ &= \bar{\sigma}\cap(Q\times Q).\end{aligned}$$

∎

Lemma 3. *Let $(P,\leq)$ be a residuated lattice, Q a subresiduated lattice of P and σ a pre-congruence on $(P,\leq)$. Then $Q^\sigma := \{p\in P \mid Q\cap[p]_{\bar{\sigma}}\neq\emptyset\}$ is a subresiduated lattice of P.*

Proof: First notice that LR_3 holds in any subset of P. Hence it is enough to show that Q^σ is closed under its binary operations. To see this, first let A be an finite subset of Q^σ. Then for all $a\in A$ there exists $q_a\in Q\cap[a]_{\bar{\sigma}}$. Take $S := \{q_a \mid a\in A\}$ and $s := \bigvee S$. So $a\ \sigma\ q_a\leq s$, for all $a\in A$. This gives that $a\ \sigma\ s$, for all $a\in A$ (notice that $\leq\subseteq\sigma$ and σ is transitive). Hence and by Theorem 1, $\bigvee A\ \sigma\ s$. Furthermore, $q_a\ \sigma\ a\leq\bigvee A$, for all $a\in A$. This gives that $q_a\ \sigma\ \bigvee A$, for all $q_a\in S$. Hence and by Theorem 1, $s=(\bigvee S)\ \sigma\ (\bigvee A)$. Consequently, $(\bigvee A)\ \sigma\ s\ \sigma\ (\bigvee A)\}$ implies that $s\in[\bigvee A]_{\bar{\sigma}}$ and so $[\bigvee A]_{\bar{\sigma}}\cap Q\neq\emptyset$ (notice that Q is a subresiduated lattice and $S\subseteq Q$). Hence $\bigvee A\in Q^\sigma$, as required.

Second, let B be a finite subset of Q^σ. Then for all $b\in B$ there exists $q_b\in Q\cap[b]_{\bar{\sigma}}$. Take $T := \{q_b \mid b\in B\}$ and $t := \bigwedge T$. So $t\leq q_b\ \sigma\ b$, for all $b\in B$. This gives that $t\ \sigma\ b$, for all $b\in B$ (notice that $\leq\subseteq\sigma$ and σ is transitive). Hence and by Theorem 1, $t\ \sigma\ \bigwedge B$. Furthermore, $\bigwedge B\leq b\ \sigma\ q_b$, for all $b\in B$. This gives that $\bigwedge B\ \sigma\ q_b$, for all $q_b\in T$. Hence and by Theorem 1, $t=(\bigwedge T)\ \sigma\ (\bigwedge B)$. Consequently, $(\bigwedge B)\ \sigma\ t\ \sigma\ (\bigwedge B)\}$ implies that $t\in[\bigwedge B]_{\bar{\sigma}}$ and so $[\bigwedge B]_{\bar{\sigma}}\cap Q\neq\emptyset$ (notice that Q is a subresiduated lattice and $S\subseteq Q$). Hence $\bigwedge B\in Q^\sigma$, as required.

Third, let $a,b\in Q^\sigma$, then there exist $q_a\in Q\cap[a]_{\bar{\sigma}}$ and $q_b\in Q\cap[b]_{\bar{\sigma}}$. So $q_a\odot q_b\in Q$ and also $q_a\ \sigma\ a$, $a\ \sigma\ q_a$ and $b\ \sigma\ q_b$, $q_b\ \sigma\ b$. By LR_2, we have $q_a\odot q_b\ \sigma\ a\odot b$ and $a\odot b\ \sigma\ q_a\odot q_b$. Therefore $q_a\odot q_b\in[a\odot b]_{\bar{\sigma}}$, this means that $[a\odot b]_{\bar{\sigma}}\cap Q\neq\emptyset$ or $a\odot b\in Q^\sigma$ as we needed.

Finally, we have $q_a\to q_b\in Q$ and $a\ \sigma\ q_a$. Then by LR_2

$$a\odot(q_a\to q_b)\ \sigma\ q_a\odot(q_a\to q_b)\ \sigma\ q_b\ \sigma\ b.$$

Hence $a\odot(q_a\to q_b)\ \sigma\ b$, so $q_a\to q_b\ \sigma\ a\to b$ by LR_3. On the other hand, from $q_a\ \sigma\ a$ we get

$$q_a\odot(a\to b)\ \sigma\ a\odot(a\to b)\ \sigma\ b\ \sigma\ q_b.$$

Then $q_a\odot a\to b\ \sigma\ q_b$, so $a\to b\ \sigma\ q_a\to q_b$. Therefore $q_a\to q_b\in[a\to b]_{\bar{\sigma}}$, this means that $[a\to b]_{\bar{\sigma}}\cap Q\neq\emptyset$ or $a\to b\in Q^\sigma$ as we needed. Consequently, Q^σ is a subresiduated lattice of P. ∎

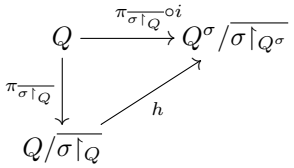
Theorem 6 (The Third Isomorphism Theorem). *Let $(P,\leq)$ be a residuated lattice, Q a subresiduated lattice of P and σ be a pre-congruence on P. Then*

$$Q/\overline{\sigma\restriction_Q}\cong Q^\sigma/\overline{\sigma\restriction_{Q^\sigma}}$$

Proof: We consider the residuated lattice homomorphism $\pi_{\overline{\sigma\restriction_{Q^\sigma}}}\circ i\colon Q\to Q^\sigma/\overline{\sigma\restriction_{Q^\sigma}}$, where $i\colon Q\to Q^\sigma$ is the inclusion map. Moreover, we have $\vec{\ker}\pi_{\overline{\sigma\restriction_{Q^\sigma}}}\circ i = \sigma\restriction_Q$. In fact,

$$\begin{aligned}(q,q')\in\vec{\ker}\pi_{\overline{\sigma\restriction_Q}}\circ i &\Leftrightarrow \pi_{\overline{\sigma\restriction_Q}}\circ i(q)\leq\pi_{\overline{\sigma\restriction_Q}}\circ i(q')\\ &\Leftrightarrow \pi_{\overline{\sigma\restriction_Q}}(q)\leq\pi_{\overline{\sigma\restriction_Q}}(q')\\ &\Leftrightarrow (q,q')\in(\vec{\ker}\pi_{\overline{\sigma\restriction_Q}})\cap(Q\times Q)\\ &\Leftrightarrow (q,q')\in(\sigma\restriction_{Q^\sigma})\cap(Q\times Q)\\ &\Leftrightarrow (q,q')\in(\sigma\cap(Q^\sigma\times Q^\sigma))\cap(Q\times Q)\\ &\Leftrightarrow (q,q')\in\sigma\cap((Q^\sigma\times Q^\sigma)\cap(Q\times Q))\\ &\Leftrightarrow (q,q')\in\sigma\cap(Q\times Q)=\sigma\restriction_Q\end{aligned}$$

So, By Corollary 1, there is a unique injective residuated lattice homomorphism $h\colon Q/\overline{\sigma\restriction_Q}\to Q^\sigma/\overline{\sigma\restriction_{Q^\sigma}}$ such that the diagram



commutative. In other words, h given by $h([q]_{\overline{\sigma\restriction_Q}}) = [q]_{\overline{\sigma\restriction_{Q^\sigma}}}$. To finish the proof, it is enough to show that h is also surjective. By the definition of Q^σ, for all $a\in Q^\sigma$ there exists a $q_a\in Q$ with $q_a\in[a]_{\bar{\sigma}}$. Thus $h([q_a]_{\overline{\sigma\restriction_Q}}) = [q_a]_{\overline{\sigma\restriction_{Q^\sigma}}} = [a]_{\overline{\sigma\restriction_{Q^\sigma}}}$, as required (notice that by Lemma 2, $\overline{\sigma\restriction_Q} = \bar{\sigma}\cap(Q\times Q)$ and $\overline{\sigma\restriction_{Q^\sigma}} = \bar{\sigma}\cap(Q^\sigma\times Q^\sigma)$). ∎

IV. Conclusions

Every function $f: A\to B$ can be factored through its image, that is, written as a composite $f = m\circ e$, where $e: A\to f(A)$ is a codomain restericton of A and $m: f(A)\to B$ is the inclusion. In other words, every function has a factorization $f = m\circ e$ where m is onto and e is one-one. So in this paper factorization structures for resituated lattice homorphisms are investigated. In [4], the authors characterized pre-congruences on a residuated lattices. Having this idea, we also proved the counterpart of algebra's first, second and third isomorphism theorems for residuated lattices.

REFERENCES

[1] P. Hajek, Mathematics of Fuzzy Logic, Dordrecht: Kluwer Academic Publshers, 1998.

[2] A. Iorgulescu, Algebras of logic as BCK algebra, Romania: Academy of Economic Studies Bucharest, 2008.

[3] P. T. Johnstone, Stone Space, Cambridge University Press 1982.

[4] F. Farsad and A. Broumand Saeid, "Residuated lattice congruences via directed kernels", submited.

[5] S. Feng and J. Yang, "Congruences in residuated lattices", 7th International Conference on Modelling, Identification and Control, 2015, pp. 18-20.

[6] M. Ward and R. P. Dilworth,"Residuated lattices", Trans. Am. Math. Soc., 45, 1939, pp. 335-354.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

The fuzzy D'Alembert solutions of the fuzzy wave equation under generalized differentiability

1st M. Miri Karbasaki, 2nd M. R. Balooch shahryari, and 3rd O. Sedaghatfar

1st Department of Mathematics, Kerman Branch, Islamic Azad University, Kerman, Iran, email address: mohsenmiri52@gmail.com
2nd Department of Mathematics, Kerman Branch, Islamic Azad University, Kerman, Iran, email address: baloochshahryari.1980@gmail.com
3rd Department of Mathematics, College of Science, Yadegar-e-Imam Khomeini (RAH) Shahre Rey Branch, Islamic Azad University, Tehran, Iran, email address: sedaghatfar@yahoo.com

***Abstract*— In this paper, we solve a one-dimensional homogeneous fuzzy wave equation with an analytical procedure using the fuzzy D'Alembert method by considering the generalized differentiability. Then, some definitions related to fuzzy numbers, theorems, and used lemmas are given. Additionally, the physical interpretation and dependency domain of fuzzy wave solutions are investigated by providing examples, where the fuzzy wave solutions are in the form of fuzzy standing, traveling, and recursive waves.**

***Index Terms*— Generalized Hukuhara differentiability; Fuzzy partial differential equation; Fuzzy wave equation; Fuzzy D'Alembert method.**

I. Introduction

Fuzzy partial differential equations often arise from the formulation of the fundamental laws of nature or mathematical analysis of uncertainty in applied mathematics and engineering sciences. Most laws of nature and physics, Newtonian motion laws, and equations such as telegraph, heat, wave, etc. are in the form of fuzzy partial differential equations in uncertainty. These laws express the phenomena of physics by connecting fuzzy space and fuzzy derivatives with time. Professor Allahviranloo et al. presented an analytical method for solving the fuzzy heat equation under generalized Hukuhara derivation (gH-derivation) [25]. In this regard, we have solved the one-dimensional fuzzy wave equation in electromagnetic and telegraph equations using an analytical technique under generalized derivatives (gH-derivatives). Besides, the physical interpretations of the ambiguous wave responses are presented by giving examples, in which the solutions are shown as the fuzzy standing wave, fuzzy traveling wave, and fuzzy backward wave. Here, we present some basic concepts of fuzzy theory, including fuzzy sets, fuzzy numbers, fuzzy new definitions, fuzzy gH-derivatives, and related concepts used in this study.

This paper aims to attain a solution for a fuzzy wave equation under generalized partial Hukuhara differentiability by the fuzzy D'Alembert method. To find the solution some properties for generalized partial Hukuhara differentiability are provided. In section II, the fuzzy wave equation model under generalized Hukuhara differentiability is defined, and the solution is studied by the fuzzy D'Alembert method. In section III, the fuzzy D'Alembert solution for the wave on a infinite string under generalized differentiability is studied. In section IV, the solution method and domain of dependence are explored. In section V, the conclusion is given, and the results are shown in some examples.

II. Fuzzy D'Alembert Solution for the Wave on a Infinite String under Generalized Hukuhara Derivative

Consider that we have a linear fuzzy partial differential equation (FPDE) in the following form

$$F\Big(u, \frac{\partial u_{gH}}{\partial t}, \frac{\partial u_{gH}}{\partial x}, \frac{\partial^2 u_{gH}}{\partial t^2}, \frac{\partial^2 u_{gH}}{\partial x^2}\Big) = \tilde{0}, \qquad (1)$$

where $u = u(x,t)$ is an unknown fuzzy function, F is a polynomial in u and its generalized Hukuhara derivatives.
With change of variable we can define the new coordinates

$$\xi(x,t) = x - ct, \quad \eta(x,t) = x + ct, \qquad (2)$$

and let

$$u(x,t) = U(\xi, \eta),$$

where $c \in \mathbb{R}^+$ is arbitrary constant generally termed the wave velocity.
With differentiating (2) $w.r.t.$ x and t yields, we have

$$\frac{\partial \xi}{\partial t} = -c, \frac{\partial \eta}{\partial t} = c, \frac{\partial \xi}{\partial x} = 1, \frac{\partial \eta}{\partial x} = 1,$$
$$\frac{\partial^2 \xi}{\partial x^2} = \frac{\partial^2 \eta}{\partial x^2} = \frac{\partial^2 \xi}{\partial t^2} = \frac{\partial^2 \eta}{\partial t^2} = 0.$$

Let us consider the homogeneous one-dimensional fuzzy wave equation

$$\text{FPDE}: \frac{\partial^2 u_{gH}(x,t)}{\partial t^2} \ominus_{gH} c^2 \odot \frac{\partial^2 u_{gH}(x,t)}{\partial x^2} = \tilde{0}, \quad -\infty < x < \infty,\ t > 0, \tag{3}$$

$$\text{FIC}_s: \begin{cases} u(x,0) = \gamma \odot f(x) = \tilde{f}(x), \\ \frac{\partial u_{gH}}{\partial t} u(x,0) = \beta \odot g(x) = \tilde{g}(x), \end{cases} \tag{4}$$

where $\gamma, \beta \in \mathbb{R}_{\mathcal{F}}$ and $f, g : \mathbb{R} \longrightarrow \mathbb{R}$, f of twice and g is once continuously differentiable and $c \in (0, +\infty)$.
Now by considering the type of $[(gH)-p]$-differentiability for $U(\xi,\eta)$, the following cases are obtained.

Case(i). Let $u(x,t)$ and $\frac{\partial u}{\partial t}$ are $[(i)-p]$-differentiable fuzzy functions $w.r.t.\ t$ then
• If $U(\xi,\eta)$ is $[(i)-gH]$-differentiable $w.r.t.\ t$ and $\frac{\partial U}{\partial \xi}$ is $[(ii)-gH]$-differentiable $w.r.t.\ t$ and $\frac{\partial U}{\partial \eta}$ is $[(i)-gH]$-differentiable $w.r.t.\ t$ fuzzy function without any switching points then

$$\begin{aligned} \frac{\partial u_{i.gH}}{\partial t} &= \frac{\partial U_{ii.gH}}{\partial \xi} \odot \frac{\partial \xi}{\partial t} \oplus \frac{\partial U_{i.gH}}{\partial \eta} \odot \frac{\partial \eta}{\partial t} \\ &= (-1)c \odot \frac{\partial U_{ii.gH}}{\partial \xi} \oplus c \odot \frac{\partial U_{i.gH}}{\partial \eta}. \end{aligned}$$

• If $U(\xi,\eta)$ is $[(i)-gH]$-differentiable $w.r.t.\ x$ and $\frac{\partial U}{\partial \xi}$ and $\frac{\partial U}{\partial \eta}$ are $[(i)-gH]$-differentiable $w.r.t.\ x$ fuzzy functions then

$$\begin{aligned} \frac{\partial u_{i.gH}}{\partial x} &= \frac{\partial U_{i.gH}}{\partial \xi} \odot \frac{\partial \xi}{\partial x} \oplus \frac{\partial U_{i.gH}}{\partial \eta} \odot \frac{\partial \eta}{\partial x} \\ &= \frac{\partial U_{i.gH}}{\partial \xi} \oplus \frac{\partial U_{i.gH}}{\partial \eta}. \end{aligned}$$

• If $U(\xi,\eta)$ is $[(i)-gH]$-differentiable $w.r.t.\ t$ and $\frac{\partial U}{\partial \xi}$ and $\frac{\partial U}{\partial \eta}$ are $[(i)-gH]$-differentiable $w.r.t.\ t$ fuzzy functions then

$$\begin{aligned} \frac{\partial^2 u_{i.gH}}{\partial t^2} &= c^2 \odot \frac{\partial^2 U_{i.gH}}{\partial \xi^2} \oplus (-1)c^2 \\ &\odot \frac{\partial}{\partial \eta}\Big(\frac{\partial U_{ii.gH}}{\partial \xi}\Big) \oplus (-1)c^2 \odot \frac{\partial}{\partial \xi}\Big(\frac{\partial U_{i.gH}}{\partial \eta}\Big) \\ &\oplus c^2 \odot \frac{\partial^2 U_{i.gH}}{\partial \eta^2}. \end{aligned} \tag{5}$$

• If $U(\xi,\eta)$ is $[(i)-gH]$-differentiable $w.r.t.\ x$ and $\frac{\partial U}{\partial \xi}$ and $\frac{\partial U}{\partial \eta}$ are $[(i)-gH]$-differentiable $w.r.t.\ x$ fuzzy functions, then

$$\begin{aligned} \frac{\partial^2 u_{i.gH}}{\partial x^2} &= \frac{\partial^2 U_{i.gH}}{\partial \xi^2} \oplus \frac{\partial}{\partial \eta}\Big(\frac{\partial U_{i.gH}}{\partial \xi}\Big) \\ &\oplus \frac{\partial}{\partial \xi}\Big(\frac{\partial U_{i.gH}}{\partial \eta}\Big) \oplus \frac{\partial^2 U_{i.gH}}{\partial \eta^2}. \end{aligned} \tag{6}$$

With substituting derivatives (5) and (6) into the FPDE (3) yields therefore , we have

$$U(\xi,\eta) = \tilde{\psi}(\xi) \oplus \tilde{\phi}(\eta).$$

Substituting for ξ and η from (2) and recalling that $u(x,t) = U(\xi,\eta)$ gives

$$u(x,t) = U(\xi,\eta) = \tilde{\psi}(x-ct) \oplus \tilde{\phi}(x+ct).$$

• If $u(x,t)$ is $[(i)-p]$-differentiable $w.r.t.\ t$, then $\tilde{\psi}$ and $\tilde{\phi}$ are $[(i)-gH]$-differentiable $w.r.t.\ (x-ct)$ and $(x+ct)$ respectively. In general, it follows that any solution to the fuzzy wave equation can be obtained as a superposition of two forward and backward waves

$$u(x,t) = \tilde{\psi}(x-ct) \oplus \tilde{\phi}(x+ct). \tag{7}$$

Now we would like to satisfy the initial conditions

$$u(x,0) = \tilde{f}(x), \tag{8}$$

$$\frac{\partial u}{\partial t}(x,0) = \tilde{g}(x). \tag{9}$$

Since equation (7) is a fuzzy solution for equation (3), then it must apply to the fuzzy initial conditions of the equation (4), hence the fuzzy initial condition $u(x,0) = \tilde{f}(x)$ concludes

$$u(x,0) = \tilde{\psi}(x) \oplus \tilde{\phi}(x) = \tilde{f}(x). \tag{10}$$

With differentiating (7) $w.r.t.\ t$ yields, we have

$$\begin{aligned} \frac{\partial u}{\partial t}(x,t) = &- c \odot \frac{\partial \tilde{\psi}}{\partial t}(x-ct) \\ &\oplus c \odot \frac{\partial \tilde{\phi}}{\partial t}(x+ct) = \tilde{g}(x), \end{aligned}$$

so that at $t = 0$ by the fuzzy initial condition (9), we obtain

$$\frac{\partial u}{\partial t}(x,0) = (-1)c \odot \tilde{\psi}'(x) \oplus c \odot \tilde{\phi}'(x) = \tilde{g}(x).$$

With dividing this last equation by c and by integration $w.r.t.\ x$ and according Definition gH-difference, we have

$$u(x,t) = \frac{1}{2}\Big(\tilde{f}(x-ct) \oplus \tilde{f}(x+ct)\Big) \oplus \frac{1}{2c}\int_{x-ct}^{x+ct} \tilde{g}(s)ds. \tag{11}$$

Suppose that $u(x,t) = u_1(x,t)$, therefore

$$u_1(x,t) = \frac{1}{2}\Big(\tilde{f}(x-ct) \oplus \tilde{f}(x+ct)\Big) \oplus \frac{1}{2c}\int_{x-ct}^{x+ct} \tilde{g}(s)ds.$$

Case(ii). Let $u(x,t)$ is $[(ii)-p]$-differentiable $w.r.t.\ t$ and $\frac{\partial u}{\partial t}$ are $[(i)-p]$-differentiable $w.r.t.\ t$ fuzzy function then,
• If $U(\xi,\eta)$ is $[(ii)-gH]$-differentiable $w.r.t.\ t$ and $\frac{\partial U}{\partial \xi}$ is $[(ii)-gH]$-differentiable $w.r.t.\ t$ and $\frac{\partial U}{\partial \eta}$ is $[(i)-gH]$-differentiable $w.r.t.\ t$ fuzzy function without any switching points then,

$$\begin{aligned} \frac{\partial u_{i.gH}}{\partial t} &= \frac{\partial U_{ii.gH}}{\partial \xi} \odot \frac{\partial \xi}{\partial t} \oplus \frac{\partial U_{ii.gH}}{\partial \eta} \odot \frac{\partial \eta}{\partial t} \\ &= (-1)c \odot \frac{\partial U_{ii.gH}}{\partial \xi} \oplus c \odot \frac{\partial U_{ii.gH}}{\partial \eta}. \end{aligned}$$

• If $U(\xi,\eta)$ is $[(ii)-gH]$-differentiable $w.r.t.\ x$ and $\frac{\partial U}{\partial \xi}$ and $\frac{\partial U}{\partial \eta}$ are $[(i)-gH]$-differentiable $w.r.t.\ x$ fuzzy functions then,

$$\frac{\partial u_{i.gH}}{\partial x} = \frac{\partial U_{i.gH}}{\partial \xi} \odot \frac{\partial \xi}{\partial x} \oplus \frac{\partial U_{i.gH}}{\partial \eta} \odot \frac{\partial \eta}{\partial x} = \frac{\partial U_{i.gH}}{\partial \xi} \oplus \frac{\partial U_{i.gH}}{\partial \eta}.$$

• If $U(\xi,\eta)$ is $[(ii)-gH]$-differentiable $w.r.t.$ t and $\frac{\partial U}{\partial \xi}$ and $\frac{\partial U}{\partial \eta}$ are $[(i)-gH]$-differentiable $w.r.t.$ t fuzzy functions then,

$$\frac{\partial^2 u_{ii.gH}}{\partial t^2} = c^2 \odot \frac{\partial^2 U_{ii.gH}}{\partial \xi^2} \oplus (-1)c^2 \odot \frac{\partial}{\partial \eta}\Big(\frac{\partial U_{ii.gH}}{\partial \xi}\Big) \oplus (-1)c^2 \odot \frac{\partial}{\partial \xi}\Big(\frac{\partial U_{ii.gH}}{\partial \eta}\Big) \oplus c^2 \odot \frac{\partial^2 U_{ii.gH}}{\partial \eta^2}. \tag{12}$$

• If $U(\xi,\eta)$ is $[(ii)-gH]$-differentiable $w.r.t.$ x and $\frac{\partial U}{\partial \xi}$ and $\frac{\partial U}{\partial \eta}$ are $[(ii)-gH]$-differentiable $w.r.t.$ x fuzzy functions then,

$$\frac{\partial^2 u_{ii.gH}}{\partial x^2} = \frac{\partial^2 U_{ii.gH}}{\partial \xi^2} \oplus \frac{\partial}{\partial \eta}\Big(\frac{\partial U_{ii.gH}}{\partial \xi}\Big) \oplus \frac{\partial}{\partial \xi}\Big(\frac{\partial U_{ii.gH}}{\partial \eta}\Big) \oplus \frac{\partial^2 U_{ii.gH}}{\partial \eta^2}. \tag{13}$$

$$u(x,t) = U(\xi,\eta) = \tilde{\psi}(x-ct) \oplus \tilde{\phi}(x+ct).$$

• $u(x,t)$ is $[(ii)-p]$-differentiable $w.r.t.$ t, then $\tilde{\psi}$ and $\tilde{\phi}$ are $[(ii)-gH]$-differentiable $w.r.t.$ $(x-ct)$ and $w.r.t.$ $(x+ct)$ respectively. In general, it follows that any solution to the fuzzy wave equation can be obtained as a superposition of two forward and backward waves

$$u(x,t) = \tilde{\psi}(x-ct) \oplus \tilde{\phi}(x+ct). \tag{14}$$

Now we would like to satisfy the initial conditions

$$u(x,0) = \tilde{f}(x), \quad \frac{\partial u}{\partial t}(x,0) = \tilde{g}(x). \tag{15}$$

Since equation (14) is a fuzzy solution for equation (3), then it must apply to the fuzzy initial conditions of the equation (4), hence the fuzzy initial condition $u(x,0) = \tilde{f}(x)$ concludes

$$u(x,0) = \tilde{\psi}(x) \oplus \tilde{\phi}(x) = \tilde{f}(x). \tag{16}$$

Differentiating from (14) $w.r.t.$ t yields, we have

$$\frac{\partial u}{\partial t}(x,t) = -\,c \odot \frac{\partial \tilde{\psi}}{\partial t}(x-ct) \oplus c \odot \frac{\partial \tilde{\phi}}{\partial t}(x+ct) = \tilde{g}(x),$$

so that at $t=0$ by fuzzy initial condition (15), we obtain

$$\frac{\partial u}{\partial t}(x,0) = (-1)c \odot \tilde{\psi}'(x) \oplus c \odot \tilde{\phi}'(x) = \tilde{g}(x).$$

With dividing this last equation by c and by integration $w.r.t.$ x and according Definition gH-difference, we have

$$u(x,t) = \frac{1}{2}\Big(\tilde{f}(x-ct) \oplus \tilde{f}(x+ct)\Big) \ominus_{gH} \frac{(-1)}{2c}\int_{x-ct}^{x+ct} \tilde{g}(x)ds. \tag{17}$$

Suppose that $u(x,t) = u_2(x,t)$, therefore

$$u_2(x,t) = \frac{1}{2}\Big(\tilde{f}(x-ct) \oplus \tilde{f}(x+ct)\Big) \ominus_{gH} \frac{(-1)}{2c}\int_{x-ct}^{x+ct} \tilde{g}(x)ds.$$

We examine solution of u_1 and u_2 according to the type gH-differentiability for the functions of $\tilde{f}, \tilde{f}'$ and $\tilde{g}, \tilde{g}'$.

Now, by using Theorem 1, Lemma 1 and Lemma 2 we show that according to gH-differentiability of $\tilde{f}$ and $\tilde{g}$ that u_1 and u_2 are the solution of FPDE 3.

Theorem 1. *Let $\tilde{f}, \tilde{g} : \mathbb{R}^2 \to \mathbb{R}_{\mathcal{F}}$ be two fuzzy functions such that $\tilde{g}$ is once continuously generalized differentiable function, $\tilde{f}$ twice and $c > 0$. Consider*

$$u_1(x,t) = \frac{1}{2}\big(\tilde{f}(x-ct) \oplus \tilde{f}(x+ct)\big) \oplus \frac{1}{2c}\int_{x-ct}^{x+ct} g(s)ds,$$

and

$$u_2(x,t) = \frac{1}{2}\big(\tilde{f}(x-ct) \oplus \tilde{f}(x+ct)\big) \ominus_{gH} \frac{(-1)}{2c}\int_{x-ct}^{x+ct} g(s)ds,$$

provided for all $(x,t) \in \mathbb{R} \times (0,\infty)$ the above gH-difference exists.
1. If $\tilde{f}, \tilde{f}'$ and $\tilde{g}, \tilde{g}'$ are $[(i)-gH]$-differentiable or $\tilde{f}, \tilde{f}'$ and $\tilde{g}, \tilde{g}'$ are $[(ii)-gH]$-differentiable on $\mathbb{R}$, then u_1 is a solution of FPDE 3.
2. If $\tilde{f}$ is $[(i)-gH]$-differentiable and $\tilde{f}'$ is $[(ii)-gH]$-differentiable and $\tilde{g}, \tilde{g}'$ are $[(i)-gH]$-differentiable (or $\tilde{g}, \tilde{g}'$ are $[(ii)-gH]$-differentiable) $w.r.t.$ x and t, then u_2 is a solution of FPDE 3.

Lemma 1. *Let $\tilde{f} : \mathbb{R}^2 \to \mathbb{R}_{\mathcal{F}}$ be a fuzzy-valued function of two variables is gH-differentiable $w.r.t.$ x, t and c is a positive constant real number for all $(x,t) \in \mathbb{R} \times (0,\infty)$. Consider*

$$\tilde{H}(x,t) = \tilde{f}(x-ct) \oplus \tilde{f}(x+ct).$$

Then
1. $\tilde{H}$ is gH-differentiable $w.r.t.$ x on $\mathbb{R}$.
2. If $\tilde{f}'$ is gH-differentiable on $\mathbb{R}$, then $\tilde{H}$ is gH-differentiable $w.r.t.$ t on $\mathbb{R} \times (0,\infty)$.

Lemma 2. *Let $\tilde{g} : \mathbb{R}^2 \to \mathbb{R}_{\mathcal{F}}$ be a fuzzy-valued function of two variables is gH-differentiable and c is a positive constant real number for $(x,t) \in \mathbb{R} \times (0,\infty)$ set*

$$\tilde{G}(x,t) = \int_{x-ct}^{x+ct} \tilde{g}(s)ds.$$

If the type of gH-derivative $\tilde{g}$ and $\tilde{g}'$ are the same $w.r.t.$ x, then the type of gH-derivative $\tilde{G}$ is also proportional to the type of gH-derivative $\tilde{g}$ and $\tilde{g}'$. Also, if the types of gH-derivative $\tilde{g}$ and $\tilde{g}'$ are different, then the gH-derivative $\tilde{G}$ is proportional to the gH-derivative of $\tilde{g}$.

Example 1. *Let us consider the homogeneous one-dimensional fuzzy wave equation*

$$\frac{\partial^2 u_{gH}}{\partial t^2}(x,t) \ominus_{gH} 9 \odot \frac{\partial^2 u_{gH}}{\partial x^2}(x,t) = \tilde{0},$$

where the fuzzy initial conditions are as follows:

$$\begin{cases} u(x,0) = \tilde{f}(x) = \langle 1,2,5\rangle \odot e^x, \quad -\infty < x < \infty, \\ \frac{\partial u_{gH}}{\partial t}(x,0) = \tilde{g}(x) = \langle 1,3,4\rangle \odot e^x. \end{cases}$$

The fuzzy solution of the wave equation u_1 is

$$\begin{aligned} u_1(x,t) &= \frac{1}{2}\Big(\gamma \odot f(x-ct) \\ &\oplus \gamma \odot f(x+ct)\Big) \oplus \frac{1}{2c}\int_{x-ct}^{x+ct} \beta \odot g(s)ds \\ &= \frac{1}{2}\Big(\langle 1,2,5\rangle \odot e^{x-3t} \oplus \langle 1,2,5\rangle \odot e^{x+3t}\Big) \\ &\oplus \frac{1}{6}\int_{x-3t}^{x+3t} \langle 1,3,4\rangle \odot e^s ds \\ &= \frac{1}{2}\langle 1,2,5\rangle \odot \Big(e^{x-3t} \oplus e^{x+3t}\Big) \oplus \frac{1}{6}\langle 1,3,4\rangle \\ &\odot \Big(e^{x+3t} \ominus_{gH} e^{x-3t}\Big). \end{aligned}$$

Also, the fuzzy solution of u_2 is

$$\begin{aligned} u_2(x,t) &= \frac{1}{2}\Big(\gamma \odot f(x-ct) \oplus \gamma \odot f(x+ct)\Big) \\ &\ominus_{gH} \frac{(-1)}{2c}\int_{x-ct}^{x+ct} \beta \odot g(s)ds \\ &= \frac{1}{2}\Big(\langle 1,2,5\rangle \odot e^{x-3t} \oplus \langle 1,2,5\rangle \odot e^{x+3t}\Big) \\ &\ominus_{gH} \frac{(-1)}{6}\int_{x-3t}^{x+3t} \langle -4,-3,-1\rangle \odot e^s ds \\ &= \frac{1}{2}\langle 1,2,5\rangle \odot \Big(e^{x-3t} \oplus e^{x+3t}\Big) \ominus_{gH} \frac{1}{6}\langle 1,3,4\rangle \\ &\odot \Big((-1)e^{x-3t} \ominus_{gH} (-1)e^{x+3t}\Big). \end{aligned}$$

III. The Fuzzy D'Alembert Solutions for the fuzzy Wave on a infinite string under generalized derivative

Let

$$\tilde{u} = \frac{\partial_g^2 \tilde{y}}{\partial t^2}(x,t), \quad \tilde{v} = c^2 \odot \frac{\partial_g^2 \tilde{y}}{\partial x^2}(x,t), \tag{18}$$

$$\tilde{w} = \frac{\partial_g^2 \tilde{y}}{\partial t^2}(x,t) \ominus_g c^2 \odot \frac{\partial_g^2 \tilde{y}}{\partial x^2}(x,t). \tag{19}$$

By using Definition (g)-difference, we have

$$\begin{aligned} &\Big[\frac{\partial_g^2 \tilde{y}}{\partial t^2}(x,t) \ominus_g c^2 \cdot \frac{\partial_g^2 \tilde{y}}{\partial x^2}(x,t)\Big]_r \\ &= \Big[\inf_{\lambda \geq r} \min\Big\{\frac{\partial^2 \underline{y}_\lambda}{\partial t^2} - c^2 \cdot \frac{\partial^2 \underline{y}_\lambda}{\partial x^2}, \frac{\partial^2 \overline{y}_\lambda}{\partial t^2} - c^2 \cdot \frac{\partial^2 \overline{y}_\lambda}{\partial x^2}\Big\} \\ &, \sup_{\lambda \geq r} \max\Big\{\frac{\partial_g^2 \underline{y}_\lambda}{\partial t^2} - c^2 \cdot \frac{\partial^2 \underline{y}_\lambda}{\partial x^2}, \frac{\partial^2 \overline{y}_\lambda}{\partial t^2} - c^2 \cdot \frac{\partial^2 \overline{y}_\lambda}{\partial x^2}\Big\}\Big] \\ &= [\underline{0}, \overline{0}]. \end{aligned} \tag{20}$$

Therefore

$$\begin{aligned} \Big[\frac{\partial_g^2 \tilde{y}}{\partial t^2}(x,t)\Big]_\lambda &= \Big[\inf_{\alpha \geq \lambda} \min\Big\{\frac{\partial^2 \underline{y}_\alpha}{\partial t^2}, \frac{\partial^2 \overline{y}_\alpha}{\partial t^2}\Big\}, \\ &\sup_{\alpha \geq \lambda} \max\Big\{\frac{\partial^2 \underline{y}_\alpha}{\partial t^2}, \frac{\partial^2 \overline{y}_\alpha}{\partial t^2}\Big\}\Big]. \end{aligned}$$

$$\begin{aligned} \Big[\frac{\partial_g^2 \tilde{y}}{\partial x^2}(x,t)\Big]_\lambda &= \Big[\inf_{\alpha \geq \lambda} \min\Big\{\frac{\partial^2 \underline{y}_\alpha}{\partial x^2}, \frac{\partial^2 \overline{y}_\alpha}{\partial x^2}\Big\}, \\ &\sup_{\alpha \geq \lambda} \max\Big\{\frac{\partial^2 \underline{y}_\alpha}{\partial x^2}, \frac{\partial^2 \overline{y}_\alpha}{\partial x^2}\Big\}\Big]. \end{aligned}$$

Thus for all $\alpha \in [0,1]$, we have

$$\frac{\partial_g^2 \underline{y}_\lambda}{\partial t^2}(x,t) = \inf_{\alpha \geq \lambda} \min\Big\{\frac{\partial^2 \underline{y}_\alpha}{\partial t^2}, \frac{\partial^2 \overline{y}_\alpha}{\partial t^2}\Big\}, \tag{21}$$

$$\frac{\partial_g^2 \overline{y}_\lambda}{\partial t^2}(x,t) = \sup_{\alpha \geq \lambda} \max\Big\{\frac{\partial^2 \underline{y}_\alpha}{\partial t^2}, \frac{\partial^2 \overline{y}_\alpha}{\partial t^2}\Big\}, \tag{22}$$

$$\frac{\partial_g^2 \underline{y}_\lambda}{\partial x^2}(x,t) = \inf_{\alpha \geq \lambda} \min\Big\{\frac{\partial^2 \underline{y}_\alpha}{\partial x^2}, \frac{\partial^2 \overline{y}_\alpha}{\partial x^2}\Big\}, \tag{23}$$

$$\frac{\partial_g^2 \overline{y}_\lambda}{\partial x^2}(x,t) = \sup_{\alpha \geq \lambda} \max\Big\{\frac{\partial^2 \underline{y}_\alpha}{\partial x^2}, \frac{\partial^2 \overline{y}_\alpha}{\partial x^2}\Big\}. \tag{24}$$

With substituting derivatives (21), (22), (23) and (24) into (20) yields

$$\begin{aligned} &\Big[\frac{\partial_g^2 \tilde{y}}{\partial t^2}(x,t) \ominus_g c^2 \odot \frac{\partial_g^2 \tilde{y}}{\partial x^2}(x,t)\Big]_r = \\ &\Big[\inf_{\lambda \geq r} \min\Big\{\inf_{\alpha \geq \lambda} \min\Big\{\frac{\partial^2 \underline{y}_\alpha}{\partial t^2} - c^2\frac{\partial^2 \underline{y}_\alpha}{\partial x^2}, \frac{\partial^2 \overline{y}_\alpha}{\partial t^2} - c^2\frac{\partial^2 \overline{y}_\alpha}{\partial x^2}\Big\}, \\ &\sup_{\lambda \geq r} \max\Big\{\sup_{\alpha \geq \lambda} \max\Big\{\frac{\partial^2 \underline{y}_\alpha}{\partial t^2} - c^2\frac{\partial^2 \underline{y}_\alpha}{\partial x^2}, \frac{\partial^2 \overline{y}_\alpha}{\partial t^2} - c^2\frac{\partial^2 \overline{y}_\alpha}{\partial x^2}\Big\}\Big] \\ &= [\underline{0}, \overline{0}]. \end{aligned}$$

Because $\alpha \geq \lambda$ and $\lambda \geq r$, then $\alpha \geq r$ therefore, we have

$$\begin{aligned} &\Big[\frac{\partial_g^2 \tilde{y}}{\partial t^2}(x,t) \ominus_g c^2 \odot \frac{\partial_g^2 \tilde{y}}{\partial x^2}(x,t)\Big]_r = \\ &\Big[\inf_{\alpha \geq r} \min\Big\{\frac{\partial^2 \underline{y}_\alpha}{\partial t^2} - c^2\frac{\partial^2 \underline{y}_\alpha}{\partial x^2}, \frac{\partial^2 \overline{y}_\alpha}{\partial t^2} - c^2\frac{\partial^2 \overline{y}_\alpha}{\partial x^2}\Big\}, \\ &\sup_{\alpha \geq r} \max\Big\{\frac{\partial^2 \underline{y}_\alpha}{\partial t^2} - c^2\frac{\partial^2 \underline{y}_\alpha}{\partial x^2}, \frac{\partial^2 \overline{y}_\alpha}{\partial t^2} - c^2\frac{\partial^2 \overline{y}_\alpha}{\partial x^2}\Big\}\Big] \\ &= [\underline{0}, \overline{0}]. \end{aligned}$$

Now, Let

$$\underline{w}_\alpha = \frac{\partial^2 \underline{y}_\alpha}{\partial t^2} - c^2\frac{\partial^2 \underline{y}_\alpha}{\partial x^2}, \qquad \overline{w}_\alpha = \frac{\partial^2 \overline{y}}{\partial t^2} - c^2\frac{\partial^2 \overline{y}}{\partial x^2}.$$

We assume the following cases:
1) If $len(\underline{w}_\alpha) \leq len(\overline{w}_\alpha)$ then

$$\inf_{\alpha\geq r} \min\left\{\underline{w}_\alpha, \overline{w}_\alpha\right\} = \inf_{\alpha\geq r} \underline{w}_\alpha = \underline{0},$$
$$\sup_{\alpha\geq r} \max\left\{\underline{w}_\alpha, \overline{w}_\alpha\right\} = \sup_{\alpha\geq r} \overline{w}_\alpha = \overline{0}.$$

2) If $len(\underline{w}_\alpha) \geq len(\overline{w}_\alpha)$ then

$$\inf_{\alpha\geq r} \min\left\{\underline{w}_\alpha, \overline{w}_\alpha\right\} = \inf_{\alpha\geq r} \overline{w}_\alpha = \underline{0},$$
$$\sup_{\alpha\geq r} \max\left\{\underline{w}_\alpha, \overline{w}_\alpha\right\} = \sup_{\alpha\geq r} \underline{w}_\alpha = \overline{0}.$$

Suppose that case 1 holds, thus we have

$$\inf_{\alpha\geq r} \underline{w}_\alpha = \inf_{\alpha\geq r}\left\{\frac{\partial^2 \underline{y}_\alpha}{\partial t^2} - c^2\frac{\partial^2 \underline{y}_\alpha}{\partial x^2}\right\} = \underline{0},$$
$$\sup_{\alpha\geq r} \overline{w}_\alpha = \sup_{\alpha\geq r}\left\{\frac{\partial^2 \overline{y}_\alpha}{\partial t^2} - c^2\frac{\partial^2 \overline{y}_\alpha}{\partial x^2}\right\} = \overline{0}.$$

On the other hand, from (11) we have

$$\underline{y}_r(x,t) = \inf_{\alpha\geq r}\Big\{\frac{1}{2}\big(\underline{\gamma}_r.f(x-ct) + \underline{\gamma}_r.f(x+ct)\big) + \frac{1}{2c}\int_{x-ct}^{x+ct} \underline{\gamma}_r.g(s)ds,\Big\}$$

and

$$\overline{y}_r(x,t) = \sup_{\alpha\geq r}\Big\{\frac{1}{2}\big(\overline{\gamma}_r.f(x-ct) + \overline{\gamma}_r.f(x+ct)\big) + \frac{1}{2c}\int_{x-ct}^{x+ct} \overline{\gamma}_r.g(s)ds\Big\}.$$

Now, suppose that case 2 holds, thus we have

$$\inf_{\alpha\geq r} \overline{w}_\alpha = \inf_{\alpha\geq r}\left\{\frac{\partial^2 \overline{y}_\alpha}{\partial t^2} - c^2\frac{\partial^2 \overline{y}_\alpha}{\partial x^2}\right\} = \underline{0},$$
$$\sup_{\alpha\geq r} \underline{w}_\alpha = \sup_{\alpha\geq r}\left\{\frac{\partial^2 \underline{y}_\alpha}{\partial t^2} - c^2\frac{\partial^2 \underline{y}_\alpha}{\partial x^2}\right\} = \overline{0},$$

Thus, we have

$$\overline{y}_r(x,t) = \inf_{\alpha\geq r}\Big\{\frac{1}{2}\big(\overline{\gamma}_r.f(x-ct) + \overline{\gamma}_r.f(x+ct)\big) + \frac{1}{2c}\int_{x-ct}^{x+ct} \overline{\gamma}_r.g(s)ds\Big\},$$

and

$$\underline{y}_r(x,t) = \sup_{\alpha\geq r}\Big\{\frac{1}{2}\big(\underline{\gamma}_r.f(x-ct) + \underline{\gamma}_r.f(x+ct)\big) + \frac{1}{2c}\int_{x-ct}^{x+ct} \underline{\gamma}_r.g(s)ds\Big\}.$$

IV. The fuzzy Solutions method and domain of dependence

At a given position $x = x_0$ on the string, the fuzzy solution in case(i) at time $t = t_0$ is

$$u(x_0,t_0) = \frac{1}{2}\Big(\tilde{f}(x_0-ct_0) \oplus \tilde{f}(x_0+ct_0)\Big) \oplus \frac{1}{2c}\int_{x_0-ct_0}^{x_0+ct_0} \tilde{g}(s)ds,$$

and the fuzzy solution in case(ii) at time $t = t_0$ is

$$u(x_0,t_0) = \frac{1}{2}\Big(\tilde{f}(x_0-ct_0) \oplus \tilde{f}(x_0+ct_0)\Big) \ominus \frac{(-1)}{2c}\int_{x_0-ct_0}^{x_0+ct_0} \tilde{g}(s)ds.$$

In other words, the fuzzy solutions are found by tracing backward and forward in time along with the characteristics $x - ct = x_0 - ct_0$ and $x + ct = x_0 + ct_0$ to the initial state $(\tilde{f}(x), \tilde{g}(x))$, then applying (11) and (17) to compute $u(x_0, t_0)$ from the initial state.
We need to determine various regions by plotting the salient characteristics $x \pm ct = const$. The regions determine where $x - ct$ and $x + ct$ are relative to the cases for the functions $\tilde{f}(x)$ and $\tilde{g}(x)$ and tells us what part of the case functions should be used in each region.

1. Assume that the case (i) is established.
For example, consider $\tilde{f}(x)$ and $\tilde{g}(x)$ of the following form

$$\tilde{f}(x) = \begin{cases}\tilde{\phi}(x), & |x| \leq l\\ 0, & |x| > l\end{cases},\ \tilde{g}(x) = \begin{cases}\tilde{\psi}(x), & |x| \leq l\\ 0, & |x| > l.\end{cases}$$

Here, we have considered zero singleton fuzzy numbers.
Step 1. Write down the fuzzy D'Alembert solution, from (11) we have

$$u(x,t) = \frac{1}{2}\Big(\tilde{f}(x-ct) \oplus \tilde{f}(x+ct)\Big) \oplus \frac{1}{2c}\int_{x-ct}^{x+ct} \tilde{g}(s)ds.$$

Step 2. Identify the regions.
The regions of interest are found the four characteristics $x \pm ct = \pm l$ and are given mathematically by

$$\begin{aligned}
R_1 &= \big\{(x,t) : -l \leq x - ct \leq l \text{ and } -l \leq x + ct \leq l\big\},\\
R_2 &= \big\{(x,t) : -l \leq x - ct \leq l \text{ and } x + ct \geq l\big\},\\
R_3 &= \big\{(x,t) : x - ct \leq -l \text{ and } -l \leq x + t \leq l\big\},\\
R_4 &= \big\{(x,t) : x - ct \leq -l \text{ and } x + ct \geq l\big\},\\
R_5 &= \big\{(x,t) : x + ct \leq -l\big\},\\
R_6 &= \big\{(x,t) : x - ct \geq l\big\}.
\end{aligned} \tag{25}$$

The regions determine where $x - ct$ and $x + ct$ are relative to $\pm l$, which tells us what part of the case functions $\tilde{f}(x)$ and $\tilde{g}(x)$ should be used. It is helpful to define the lines.
$x_A(t) = -l - ct$, $x_B(t) = -l + ct$, $x_C(t) = l - ct$, $x_D(t) = l + ct$.
Step 3. Determine the fuzzy solution in each region.

$$u(x,t)=\begin{cases}\frac{1}{2}\Big(\tilde{\phi}(x-ct)\oplus\tilde{\phi}(x+ct)\Big)\\ \oplus\frac{1}{2c}\int_{x-ct}^{x+ct}\tilde{\psi}(s)ds, & (x,t)\in R_1,\\ \frac{\tilde{\phi}(x-ct)}{2}\oplus\frac{1}{2}\int_{x-ct}^{l}\tilde{\psi}(s)ds, & (x,t)\in R_2,\\ \frac{\tilde{\phi}(x+ct)}{2}\oplus\frac{1}{2}\int_{-l}^{x+ct}\tilde{\psi}(s)ds, & (x,t)\in R_3,\\ \frac{1}{2}\int_{-l}^{l}\tilde{\psi}(s)ds, & (x,t)\in R_4,\\ 0, & (x,t)\in R_5,R_6.\end{cases}\tag{26}$$

Step 4. For each specific time t_0, write the x-intervals corresponding to the intersection of the sets R_n with the line $t=t_0$. At time 0, we use Table I to find the x intervals R_n

TABLE I
THIS AMOUNTS TO COMPUTING THE VALUES OF $x_A(t), x_B(t), x_C(t)$ AND $x_D(t)$ FOR EACH TIME

t	0	$\frac{1}{2}$	1	2
$x_A(t)$	$-l$	$-l-\frac{1}{2}c$	$-l-c$	$-l-2c$
$x_B(t)$	$-l$	$-l+\frac{1}{2}c$	$-l+c$	$-l+2c$
$x_C(t)$	l	$l-\frac{1}{2}c$	$l-c$	$l-2c$
$x_D(t)$	l	$l+\frac{1}{2}c$	$l+c$	$l+2c$

corresponding to the intersection of R_n' with the line $t=0$

$$R_5'=(-\infty,-l],R_1'=(-l,l],R_6'=[l,\infty).$$

In region R_1, we have (recall that $t=0$)

$$\begin{aligned}u(x,t)=&\frac{1}{2}\Big(\tilde{\phi}(x-0)\oplus\tilde{\phi}(x+0)\Big)\\&\oplus\frac{1}{2c}\int_{x-0}^{x+0}\tilde{\psi}(s)ds=\tilde{\phi}(x).\end{aligned}$$

Similarly, we can check that in the other regions, $u=0$, so that (26) becomes

$$\begin{aligned}u(x,0)&=\begin{cases}\tilde{\phi}(x), & x\in R_1'=[-l,l]\\ 0, & x\in R_5'\cup R_6'\end{cases}\\&=\begin{cases}\tilde{\phi}(x), & |x|\le l\\ 0, & |x|>l\end{cases}=\tilde{f}(x).\end{aligned}$$

At time $1/2$, we use Table I to find the x intervals R_n' corresponding to the intersection of R_n with the line $t=\frac{1}{2}$

$$\begin{aligned}&R_5=\big(-\infty,-l-\frac{1}{2}c\big],R_1=\big(-l+\frac{1}{2}c,l-ct\big],\\&R_6=\big[l+\frac{1}{2}c,\infty\big),R_3=\big(-l-\frac{1}{2}c,-l+\frac{1}{2}c\big],\\&R_2=\big(l-\frac{1}{2}c,l+\frac{1}{2}c\big].\end{aligned}\tag{27}$$

At time 1, we use Table I to find the x intervals R_n' corresponding to the intersection of R_n with the line $t=1$

$$\begin{aligned}&R_5=\big(-\infty,-l-c\big],R_3=\big(-l-c,-l+c\big],\\&R_2=\big[l-c,l+c\big],R_6=\big[l+c,\infty\big].\end{aligned}\tag{28}$$

At time 2, we use Table I to find the x intervals R_n' to the intersection of R_n with the line $t=2$

$$\begin{aligned}&R_5=\big(-\infty,-l-2c\big],R_3=\big[-l-2c,-l+2c\big],\\&R_4=\big[-l+2c,l-2c\big],R_2=\big[l-2c,l+2c\big],\\&R_6=\big[l+2c,\infty\big).\end{aligned}\tag{29}$$

Here, to illustrate the ability and reliability of the aforementioned concepts we have solved some application examples.

Example 2. *For an infinitely long string, consider giving the string zero fuzzy initial displacement $u(x,0)=0$ and the fuzzy initial velocity $\partial_t u(x,0)=\tilde{g}(x)$. Suppose that*

$$\tilde{g}(x)=\begin{cases}\gamma\odot cos^2(\frac{\pi}{2}x), & -1\le x\le 1,\\ 0, & otherwise.\end{cases}$$

The FICs have the form considered above for $\tilde{\phi}(x)=0$ and $\tilde{\psi}(x)=\gamma\odot cos^2(\frac{\pi}{2}x)$.

Step 1. *The fuzzy D'Alembert solution of* (11) *becomes*

$$u(x,t)=\frac{1}{2}\int_{x-t}^{x+t}\tilde{g}(s)ds.$$

Step 2. *By* (25) *we determine the regions.*

Step 3. *Determine $u(x,t)$ in each region. From* (27) *by consider $c=1,l=1,\beta=\langle 1,3,5\rangle$, we have*

$$u(x,t)=\begin{cases}\frac{1}{2}\int_{x-t}^{x+t}\langle 1,3,5\rangle\odot\psi(s)ds, & (x,t)\in R_1,\\ \frac{1}{2}\int_{x-t}^{1}\langle 1,3,5\rangle\odot\psi(s)ds, & (x,t)\in R_2,\\ \frac{1}{2}\int_{-1}^{x+t}\langle 1,3,5\rangle\odot\psi(s)ds, & (x,t)\in R_3,\\ \frac{1}{2}\int_{-1}^{1}\langle 1,3,5\rangle\odot\psi(s)ds, & (x,t)\in R_4,\\ 0, & (x,t)\in R_5,R_6.\end{cases}$$

Thus

$$u(x,t)=\begin{cases}\langle 1,3,5\rangle\odot\Big(\frac{t}{2}\oplus\frac{1}{4\pi}sin(\pi(x+t))\ominus_{gH}\\ \frac{1}{4\pi}sin(\pi(x-t))\Big),\\ \qquad\qquad (x,t)\in R_1,\\ \langle 1,3,5\rangle\odot\Big(\frac{1}{4}\ominus_{gH}\frac{1}{4}(x-t)\Big)\ominus_{gH}\\ \Big(\frac{1}{4\pi}sin(\pi(x-t)\Big),\\ \qquad\qquad (x,t)\in R_2,\\ \langle 1,3,5\rangle\odot\Big(\frac{1}{4}(x+t)\ominus_{gH}\frac{1}{4}\Big)\\ \oplus\frac{1}{4\pi}sin(\pi(x+t)\Big), & (x,t)\in R_3,\\ \langle 1,3,5\rangle\odot\frac{1}{2}, & (x,t)\in R_4,\\ 0, & (x,t)\in R_5,R_6.\end{cases}\tag{30}$$

Step 4. *We early consider, intermediate and later times, $t = 1/2, 1, 2$. At $t = 1/2$, the regions R_n are given by* (27) *and* (30) *becomes*

$$u(x,\frac{1}{2}) = \begin{cases} \langle 1,3,5\rangle \odot \Big(\frac{t}{4} \oplus \frac{1}{2\pi} cos\pi x\Big), \\ \qquad\qquad\qquad\qquad -\frac{1}{2} \le x \le \frac{1}{2}, \\ \langle 1,3,5\rangle \odot \Big(\frac{1}{4} \ominus_{gH} \frac{1}{4}(x-\frac{1}{2}) \oplus \frac{cos\pi x}{4\pi}\Big), \\ \qquad\qquad\qquad\qquad \frac{1}{2} \le x \le \frac{3}{2} \\ \langle 1,3,5\rangle \odot \Big(\frac{1}{4}(x+t) \ominus_{gH} \frac{1}{4}\Big) \\ \oplus \frac{1}{4\pi} sin(\pi(x+\frac{1}{2})\Big), \qquad -\frac{3}{2} \le x \le -\frac{1}{2}, \\ 0, \qquad\qquad\qquad\qquad x \ge \frac{3}{2}. \end{cases}$$

At time 1*, the regions R_n are given by* (28) *and* (30) *becomes*

$$u(x,1) = \begin{cases} \langle 1,3,5\rangle \odot \Big(\frac{1}{2} \ominus_{gH} \frac{1}{4}\Big) \ominus_{gH} \frac{1}{4\pi} sin\pi x, \\ \qquad\qquad\qquad\qquad 0 \le x \le 2, \\ \langle 1,3,5\rangle \odot \big(\frac{1}{4}x - \frac{1}{4\pi} sin\pi x\big), \\ \qquad\qquad\qquad\qquad -2 \le x \le 0, \\ 0, \qquad\qquad\qquad\qquad |x| \ge 2. \end{cases}$$

At time 2*, the regions R_n are given by* (29) *and* (30) *becomes*

$$u(x,2) = \begin{cases} \langle 1,3,5\rangle \odot \Big(\frac{1}{4} \ominus_{gH} \frac{1}{4}(x-2)\Big) \ominus_{gH} sin\pi x, \\ \qquad\qquad\qquad\qquad 1 \le x \le 3, \\ \langle 1,3,5\rangle \odot \Big(\frac{1}{4}(x+2) \ominus_{gH} \frac{1}{4}\Big) \oplus \frac{1}{4\pi} sin\pi x, \\ \qquad\qquad\qquad\qquad -3 \le x \le -1, \\ \frac{1}{2} \odot \langle 1,3,5\rangle, \\ \qquad\qquad\qquad\qquad -1 \le x \le 1, \\ 0, \qquad\qquad\qquad\qquad |x| \ge 3. \end{cases}$$

2. *Assume that the case (ii) is established.*
For example, consider $\tilde{f}(x)$ and $\tilde{g}(x)$ of the following form

$$\tilde{f}(x) = \begin{cases} \tilde{\phi}(x), & |x| \le l, \\ 0, & |x| > l, \end{cases} \quad \tilde{g}(x) = \begin{cases} \tilde{\psi}(x), & |x| \le l, \\ 0, & |x| > l. \end{cases}$$

Step 1. *Write down the fuzzy D'Alembert solution*

$$u(x,t) = \frac{1}{2}\Big(\tilde{f}(x-ct) \oplus \tilde{f}(x+ct)\Big) \ominus_{gH} \frac{(-1)}{2c} \int_{x-ct}^{x+ct} \tilde{g}(s)ds.$$

Step 2. *Identify the regions. (similarly, as in the previous)*

Step 3. *Determine the fuzzy solution of in each region.*

$$u(x,t) = \begin{cases} \frac{1}{2}\Big(\tilde{\phi}(x-ct) \oplus \tilde{\phi}(x+ct)\Big) \\ \ominus_{gH} \frac{(-1)}{2c} \int_{x-ct}^{x+ct} \tilde{\psi}(s)ds, \qquad (x,t) \in R_1, \\ \frac{\tilde{\phi}(x-ct)}{2} \ominus_{gH} \frac{(-1)}{2} \int_{x-ct}^{l} \tilde{\psi}(s)ds, \\ \qquad\qquad\qquad\qquad (x,t) \in R_2, \\ \frac{\tilde{\phi}(x+ct)}{2} \ominus_{gH} \frac{(-1)}{2} \int_{-l}^{x+ct} \tilde{\psi}(s)ds, \\ \qquad\qquad\qquad\qquad (x,t) \in R_3, \\ \ominus_{gH} \frac{(-1)}{2} \int_{-l}^{l} \tilde{\psi}(s)ds, \qquad (x,t) \in R_4, \\ 0, \qquad\qquad\qquad\qquad (x,t) \in R_5, R_6. \end{cases} \tag{31}$$

Step 4. *For each specific time t_0, write the x-intervals corresponding to the intersection of the sets R_n with the line $t = t_0$. (similarly, as in the previous step 4)*

Example 3. *For an infinitely long string, consider giving the string zero fuzzy initial displacement $u(x,0) = 0$ and the fuzzy initial velocity $\partial_t u(x,0) = \tilde{g}(x)$. Suppose that*

$$\tilde{g}(x) = \begin{cases} \gamma \odot cos^2(\frac{\pi}{2}x), & -1 \le x \le 1, \\ 0, & otherwise. \end{cases}$$

The FICs have the form considered above for $\tilde{\phi}(x) = 0$ and $\tilde{\psi}(x) = \gamma \odot cos^2(\frac{\pi}{2}x)$.

$$u(x,t) = \begin{cases} \ominus_{gH}\langle 1,3,5\rangle \odot \Big(\frac{(-1)}{2}t \oplus \frac{1}{4\pi} sin(\pi(x-t)) \\ \ominus_{gH} \frac{1}{4\pi} sin(\pi(x+t))\Big), \qquad (x,t) \in R_1, \\ \ominus_{gH}\langle 1,3,5\rangle \odot \Big(\frac{1}{4}(x-t) \ominus_{gH} \\ \frac{1}{4} \oplus \frac{1}{4\pi} sin(\pi(x-t)\Big), \qquad (x,t) \in R_2, \\ \ominus_{gH}\langle 1,3,5\rangle \odot \Big(\frac{1}{4} \ominus_{gH} \frac{1}{4}(x+t)\Big) \ominus_{gH} \\ \frac{(-1)}{4\pi} sin(\pi(x+t)\Big), \qquad (x,t) \in R_3, \\ \ominus_{gH}\langle 1,3,5\rangle \odot \frac{(-1)}{2} \qquad , \qquad (x,t) \in R_4, \\ 0, \qquad\qquad\qquad (x,t) \in R_5, R_6. \end{cases} \tag{32}$$

Step 4. *We early consider, intermediate and later times, $t = 1/2, 1, 2$. At $t = 1/2$, the regions R_n are given by* (27) *and* (32) *becomes*

$$u(x,\frac{1}{2}) = \begin{cases} \ominus_{gH}\langle 1,3,5\rangle \odot \Big(\frac{(-1)}{4} - \frac{1}{2\pi} cos\pi x\Big), \\ \qquad\qquad\qquad\qquad -\frac{1}{2} \le x \le \frac{1}{2}, \\ \ominus_{gH}\langle 1,3,5\rangle \odot \Big(\frac{1}{4}(x-\frac{1}{2}) \ominus_{gH} \frac{1}{4} \\ \oplus \frac{1}{4\pi} cos\pi x\Big), \qquad\qquad \frac{1}{2} \le x \le \frac{3}{2} \\ \ominus_{gH}\langle 1,3,5\rangle \odot \Big(\frac{1}{4} \ominus_{gH} \frac{1}{4}(x+\frac{1}{2})\Big) \\ \ominus_{gH} \frac{(-1)}{4\pi} cos\pi x\Big), \qquad -\frac{3}{2} \le x \le -\frac{1}{2}, \\ 0, \qquad\qquad x \ge \frac{3}{2}. \end{cases}$$

At time 1*, the regions* R_n *are given by* (28) *and* (32) *becomes*

$$u(x,1)=\begin{cases}\ominus_{gH}\langle 1,3,5\rangle\odot\Big(\frac{1}{4}(x-1)\ominus_{gH}\frac{1}{4}\Big)\\ \oplus\frac{(-1)}{4\pi}sin\pi x, & 0\le x\le 2,\\ \ominus_{gH}\langle 1,3,5\rangle\odot\big(\frac{1}{4}\ominus_{gH}\frac{1}{4}(x+1)\\ \ominus_{gH}\frac{1}{4\pi}sin\pi x\big), & -2\le x\le 0,\\ 0, & |x|\ge 2.\end{cases}$$

At time 2*, the regions* R_n *are given by* (29) *and* (32) *becomes*

$$u(x,2)=\begin{cases}\ominus_{gH}\langle 1,3,5\rangle\odot\Big(\frac{1}{4}(x-2)\ominus_{gH}\frac{1}{4}\oplus\frac{1}{4\pi}\\ sin\pi x\Big), & 1\le x\le 3,\\ \ominus_{gH}\langle 1,3,5\rangle\odot\Big(\frac{1}{4}\ominus_{gH}\frac{1}{4}(x+2)\ominus_{gH}\frac{(-1)}{4\pi}\\ sin\pi x\Big), & -3\le x\le -1,\\ \ominus_{gH}\langle 1,3,5\rangle\odot\frac{(-1)}{2}, & -1\le x\le 1,\\ 0, & |x|\ge 3.\end{cases}$$

V. Conclusion

The physics phenomena can be expressed by connecting fuzzy space and fuzzy derivatives with time, such as wave equations in electromagnetic. In this paper, the one-dimensional fuzzy wave and telegraph equations have been solved by an analytical technique under generalized Hukuhara derivatives. The basic concepts of the fuzzy theory have been presented. Considering physical interpretations, we have discussed fuzzy wave solutions obtained in detail using position variable restriction. Then, the fuzzy wave solutions have been clarified by providing examples using the fuzzy D'Alembert method.

References

[1] A. Mohebbi, and M. Dehghan. "High order compact solution of the one-space-dimensional linear hyperbolic equation." Numerical Methods for Partial Differential Equations: An International Journal 24, no. 5 (2008): 1222-1235.

[2] A. Jeffrey, Applied partial differential equations: an introduction. Academic Press, (2003).

[3] A. Jeffrey, Advanced engineering mathematics. Elsevier, (2001).

[4] B. Bede, Mathematics of fuzzy sets and fuzzy logic, Springer, London, (2013).

[5] B. Bede, S. G. Gal, Generalizations of the differentiability of fuzzy-number-valued functions with applications to fuzzy differential equations, Fuzzy Sets Syst. 151 (3) (2005) 581-599.

[6] B. Bede, L. Stefanini, Generalized differentiability of fuzzy-valued functions, Fuzzy Sets Syst. 230 (2013) 119-141.

[7] B. Bede, L. Stefanini, Generalized differentiability of fuzzy-valued functions, Fuzzy Sets Syst (2012), http://dx.doi.org/10.1016/j.fss 2012.10.003.

[8] D. Dubois, and P. Henri. "Towards fuzzy differential calculus part 3: Differentiation." Fuzzy sets and systems 8, no. 3 (1982): 225-233.

[9] D. M. Pozar, A modern course in microwave engineering IEEE Transactions on Education 33, no. 1 (1990): 129-134.

[10] D. J. Evans, H. and Bulut, The numerical solution of the telegraph equation by the alternating group explicit (AGE) method. International journal of computer mathematics, 80 (10) (2003) pp.1289-1297.

[11] G. Böhme, Non-Newtonian fluid mechanics. Vol. 31. Elsevier, (2012).

[12] L. A. Zadeh, Fuzzy sets, Inform. Control. 8 (1965) 338-353.

[13] L. Stefanini, A generalization of Hukuhara difference and division for interval and fuzzy arithmetic, Fuzzy Sets Syst. 161 (2010) 1564-1584.

[14] L. Stefanini, B. Bede, Generalized Hukuhara differentiability of interval-valued functions and interval differential equations, Nonlinear Anal., Theory Methods Appl. 71 (34) (2009) 1311-1328.

[15] M. Puri, and D. Ralescu, Differentials of functions, J. Math. Anal. Appl., 91(1983), 552-558.

[16] O. Kaleva, Fuzzy differential equations, Fuzzy Sets and Systems, 24 (1987), 301-317.

[17] P. M. Jordan, M. R. Meyer, and A. Puri, Causal implications of viscous damping in compressible fluid flows. Physical Review E, 62 (6) (2000) p.7918.

[18] Q. Liu, F. Liu, I. Turner, V. Anh, and Y. T. Gu, A RBF meshless approach for modeling a fractal mobile/immobile transport model. Applied Mathematics and Computation, 226 (2014) pp.336-347.

[19] R. Alikhani, F. Bahrami, S. Parvizi, Differential calculus of fuzzy multi-variable functions and its applications to fuzzy partial differential equations, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.011.

[20] R. Goetschel, and W. Voxman, Elementrary fuzzy calculus, fuzzy sets and systems, 18(1986), 31-43.

[21] R. Ghasemi Moghaddam, S. Abbasbandy and M. Rostamy-Malkhalifeh, A Study on Analytical Solutions of the Fuzzy Partial Differential Equations, Int. J. Industrial Mathematics, 4, (2020).

[22] R. G. Moghaddam, T. Allahviranloo, On the fuzzy Poisson equation, Fuzzy Sets and Systems 347 (2018) 105-128.

[23] R. K. Mohanty, New unconditionally stable difference schemes for the solution of multi-dimensional telegraphic equations. International Journal of Computer Mathematics 86, no. 12 (2009): 2061-2071.

[24] S. S. L. Chang, and L. A. Zadeh. "On fuzzy mapping and control." In Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers by Lotfi A Zadeh, pp. 180-184. (1996).

[25] T. Allahviranloo et al., On fuzzy solutions for heat equation based on generalized Hukuhara differentiability, Fuzzy Sets and Systems (2014).

[26] T. Allahviranloo, Z. Gouyandeh, A. Armand, A. Hasanoglu, On fuzzy solutions for heat equation based on generalized Hukuhara differentiability, Fuzzy Sets and Systems 265 (2015) 1-23.

[27] V. Lakshmikantham, T. Bhaskar, J. Devi, Theory of Set Differential Equations in Metric Spaces, Cambridge

Scientific Publishers, (2006).
[28] W. Congxin, M. Ming, Embedding problem of fuzzy number space. Part III, Fuzzy Sets Syst. 46 (2) (1992) 281- 286.
[29] Y. Chalco-Cano, R. Rodriguez-Lopez, M.D. Jimenez-Gamero, Characterizations of generalized differentiable fuzzy function, Fuzzy Sets and Systems(2015).

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

The new modifications of distance measures on hesitant fuzzy numbers

Abazar Keikha

Department of Mathematics and Statistics, *Velayat University*, Iranshahr, IRAN, abazar_keikha@yahoo.com

***Abstract*— Uncertainty as an important factor in modeling and solving practical problems, absorbed many attentions, and then many tools have been proposed to model it scientifically. Hesitant fuzzy sets are one of these tools which are suitable when the decision maker is able to expresses her/his hesitation degrees by a finite set of some values in $[0, 1]$. Hesitant fuzzy numbers are special case of hesitant fuzzy sets that have been used to solve organized complexity problems. In this paper, new types of three famous distance functions, i.e., Hamming distance, Euclidean distance, and Generalized distance functions will be defined. Then, a numerical is used to shows the impact of the new concepts comparatively with previous ones.**

***Index Terms*—Hesitant fuzzy sets, Hesitant fuzzy numbers, Hamming distance, Euclidean distance, Generalized distance**

I. Introduction

Uncertainty which had been considered as an unscientific aspect and must be removed, accepted as an important aspect of solving real world' problems, now [10]. So, many tools such as Dempster-Shafer and Probability theory, imprecise probability theory, rough sets theory, fuzzy sets & Possibility theory, have been presented to deal with it [3]. Fuzzy sets (FSs) theory [18], in which a single value from $[0, 1]$ is used to each element of FS to express its membership degree to the given FS, get more attention and developed in both theoriticaly and applicability. FSs converted to Type-2 Fuzzy sets, if membership degrees are also FSs [5]. If the elements of FSs are labeled by two values, as membership and non-membership degrees, from $[0, 1]$, we have intuitionistic fuzzy sets (IFSs) [1]. Pythagorean fuzzy sets (PFSs) [16], rough sets and its merging with IFSs [2], hesitant fuzzy sets (HFSs) [15], and intuitionistic hesitant fuzzy sets (IHFSs) [13] are latest extensions of FSs, which are proposed to model uncertainty.
The problems in the real world were have long been divided into organized simplicity (OS) problems, organized complexity (OC) problems, and disorganized complexity (DC) problems [17]. OC problems containing of finite number of factors, while the others have at most four (OS problems) or infinite factors (DC problems). From this point of view, HFSs are more suitable to model OC problems, and as a more useful method of quantifying uncertainty, soon absorbed the attention of many researchers to solve practical problems such as decision-making problems [12].
Hesitant fuzzy numbers (HFNs) are the special case of HFSs which are containing a real value together with a hesitant fuzzy element (HFEs) as membership part [4], [6]–[8]. HFNs can be used when we have a given crisp value which is not accepted/rejected completely by decision makers (DMs), and their hesitation degrees about it are given by a HFE.

II. Priliminaries

The mathematical representation of a HFS is $E = \{< x, h(x) > |x \in X\}$, where $h(x) = \{\eta_1, \eta_2, \cdots, \eta_n\}$, $\eta_i \in [0, 1]$, called HFEs [15]. For simplicity HFEs are applied in real applications of HFSs and scientific researches. Operational laws and calculus of HFSs have been defined based on triangular norm (t-norm) and triangular conorm (t-conorm) $f : [0, 1] \times [0, 1] \to [0, 1]$, with boundary condition, commutativity, associativity, and monotonicity properties [11].

Definition 1: Let $T : [0, 1] \times [0, 1] \to [0, 1]$ and $S : [0, 1] \times [0, 1] \to [0, 1]$ be two arbitrary functions. They are called t-norm and t-conorm, respectively, if $\forall x, y, z \in [0, 1]$:

$$(I)T(1, x) = x,$$
$$(II)T(x, y) = T(y, x),$$
$$(III)T(x, T(y, z)) = T(T(x, y), z),$$
$$(IV)x \leq x' \& y \leq y' \Rightarrow T(x, y) \leq T(x', y').$$

$$(I')S(0, x) = x,$$
$$(II')S(x, y) = S(y, x),$$
$$(III')S(x, S(y, z)) = S(S(x, y), z),$$
$$(IV')x \leq x' \& y \leq y' \Rightarrow S(x, y) \leq S(x', y').$$

If for all $x \in [0, 1] : T(x, x) < x$, and $S(x, x) > x$, then we have archimedean t-norm T and archimedean t-conorm S. Consider an additive generator $g : [0, 1] \to [0, +\infty)$, and $f(t) = g(1 - t)$. Then, archimedean t-norm $T(x, y) = g^{-1}(g(x) + g(y))$, and archimedean t-conorm

$S(x,y) = f^{-1}(f(x)+f(y))$, which are strictly increasing, are called strictly archimedean t-norm T and t-conorm S, respectively [9]. There are different types of T and S, based on additive generator g [11]: $g(t) = -\log t$ gives us Algebraic t-norm and t-conorm; $g(t) = \log \frac{2-t}{t}$ resulted Einstein t-norm and t-conorm; $g(t) = \log \frac{\nu+(1-\nu)t}{t},\ \nu > 0$ gives Hamacher t-norm and t-conorm.

Using these concepts, arithmetic operations of any arbitrary HFEs $h,\ h_1$, and h_2, would be defined as follows [19]:

$$\begin{aligned}
(1)\ h_1 \oplus h_2 &= \bigcup_{\eta_1\in h_1, \eta_2 \in h_2} \big\{S(\eta_1,\eta_2)\big\} \\
&= \bigcup_{\eta_1\in h_1, \eta_2 \in h_2} \big\{f^{-1}(f(\eta_1)+f(\eta_2))\big\}; \\
(2)\ h_1 \otimes h_2 &= \bigcup_{\eta_1\in h_1, \eta_2 \in h_2} \big\{T(\eta_1,\eta_2)\big\} \\
&= \bigcup_{\eta_1\in h_1, \eta_2 \in h_2} \big\{g^{-1}(g(\eta_1)+g(\eta_2))\big\}; \\
(3)\ \lambda h &= \bigcup_{\eta\in h} \big\{f^{-1}(\lambda f(\eta))\big\}, \quad \lambda > 0; \\
(4)\ h^{\lambda} &= \bigcup_{\eta\in h} \big\{g^{-1}(\lambda g(\eta))\big\}, \quad \lambda > 0;
\end{aligned}$$

where, λ is a positive real value.

Two HFEs h_1, and h_2 are called adjusted HFEs, if they are equal in cardinality, i.e. $|h_1| = |h_2|$.

In some practical decision making problems, there exist some predetermined values which are not fully accepted by decision makers (DMs) and must be used in decision making process. HFNs, which are able to simultaniously apply the given crisp values and opinions of DMs, have been proposed to model such situations.

Definition 2: Suppose a be a predetermined positive real value about an element x of the reference set X, which is hesitated by HFE $h(a) = \{\eta_1, \eta_2, \ldots, \eta_n\}, \eta_i \in [0,1]$. A HFN $\tilde{a}_H$ with mathematical representation $\langle a; h(a)\rangle$ merges these two separate information on element $x \in X$, i.e. $\tilde{a}_H = \langle a; \{\eta_1, \eta_2, \ldots, \eta_n\}\rangle$.

It means that HFNs have two parts: real part (is a predetermined crisp value), and membership part (is a HFE containing the hesitation degrees of DMs about the real part or direct assessment). HFNs are called adjusted, if their membership parts are adjusted HFEs.

Definition 3: Suppose $\tilde{a}_H = \langle a; h(a)\rangle$, $\tilde{b}_H = \langle b; h(b)\rangle$ be two adjusted HFNs. Then

$$\begin{aligned}
(1)\ \tilde{a}_H \oplus \tilde{b}_H &= \left\langle a+b; \bigcup_{\substack{\eta_{1(i)}\in h(a),\\ \eta_{2(i)}\in h(b)}} \big\{S(\eta_{1(i)}, \eta_{2(i)})\big\} \right\rangle = \\
&\left\langle a+b; \bigcup_{\substack{\eta_{1(i)}\in h(a),\\ \eta_{2(i)}\in h(b)}} \big\{f^{-1}(f(\eta_{1(i)}) + f(\eta_{2(i)}))\big\} \right\rangle, \\
(2)\ \tilde{a}_H \otimes \tilde{b}_H &= \left\langle a.b; \bigcup_{\substack{\eta_{1(i)}\in h(a),\\ \eta_{2(i)}\in h(b)}} \big\{T(\eta_{1(i)}, \eta_{2(i)})\big\} \right\rangle = \\
&\left\langle a.b; \bigcup_{\substack{\eta_{1(i)}\in h(a),\\ \eta_{2(i)}\in h_(b)}} \big\{g^{-1}(g(\eta_{1(i)}) + g(\eta_{2(i)}))\big\} \right\rangle, \\
(3)\ \lambda\tilde{a}_H &= \left\langle \lambda a; \bigcup_{\eta\in h(a)} \big\{f^{-1}(\lambda f(\eta))\big\} \right\rangle, \\
(4)\ (\tilde{a}_H)^{\lambda} &= \left\langle a^{\lambda}; \bigcup_{\eta\in h(a)} \big\{g^{-1}(\lambda g(\eta))\big\} \right\rangle,
\end{aligned}$$

where, $\lambda \in \mathbb{R}^{\geq 0}$, $\{\eta_{l(1)}, \eta_{l(2)}, \cdots\}$ is a permutation of $\{\eta_{l1}, \eta_{l2}, \cdots\}$ such that $\eta_{l(1)} \leq \eta_{l(2)} \leq \cdots$.

For some special case of t-norm and t-conorm, we can define the arithmetic operations of HFNs.

Definition 4: Let $w = (w_1, w_2, \cdots, w_k)$ with $w_i \in [0,1]$ and $\sum_{i=1}^{k} w_i = 1$ be the weight vector of HFNs $\tilde{a}_H^i = \langle a_i, h(a_i)\rangle,\ i = 1, 2, \cdots k$. Then

- $HWAA_w(\tilde{a}_H^1, \tilde{a}_H^2, \cdots, \tilde{a}_H^k) = \left\langle \sum_{i=1}^{k} w_i a_i; \bigcup_{i=1}^{k} h(a_i) \right\rangle$,

  is called hesitant fuzzy weighted arithmetic average (HWAA) operator where,

  $$\bigcup_{i=1}^{k} h(a_i) = \bigcup_{\eta_i\in h(a_i)} \max\{\eta_1, \eta_2, \cdots, \eta_k\}.$$

  The hesitant fuzzy weighted arithmetic average operator called hesitant fuzzy arithmetic average (HAA) operator if $w = \left(\frac{1}{n}, \frac{1}{n}, \cdots, \frac{1}{n}\right)$.
- $HWGA_w(\tilde{a}_H^1, \tilde{a}_H^2, \cdots, \tilde{a}_H^k) = \left\langle \prod_{i=1}^{k} a_i^{w_i}; \bigcap_{i=1}^{k} h(a_i) \right\rangle$,

  is called hesitant fuzzy weighted geometric average (HWGA) operator where,

  $$\bigcap_{i=1}^{k} h(a_i) = \bigcup_{\eta_i\in h(a_i)} \min\{\eta_1, \eta_, \cdots, \eta_k\}.$$

  The hesitant fuzzy weighted geometric average operator called hesitant fuzzy geometric average (HGA) operator, if $w = \left(\frac{1}{n}, \frac{1}{n}, \cdots, \frac{1}{n}\right)$.

Computing the distances between HFNs is an important concept, that is defined as follows.

Definition 5: Let $\tilde{a}_H = \langle a; \{\eta_1, \eta_2, \ldots, \eta_m\}\rangle$, and $\tilde{b}_H = \langle b; \{\lambda_1, \lambda_2, \ldots, \lambda_m\}\rangle$ be two arbitrary given adjusted HFNs. Then, their Generalized hesitant distance $d_G(\tilde{a}_H, \tilde{b}_H)$, Hamming hesitant distance $d_H(\tilde{a}_H, \tilde{b}_H)$, and Euclidean hesitant distance $d_E(\tilde{a}_H, \tilde{b}_H)$ will be defined as follows:

$$d_G(\tilde{a}_H, \tilde{b}_H) = \big[|a-b|^\lambda + \tfrac{1}{m}\sum_{j=1}^{m} |\eta_{(j)} - \lambda_{(j)}|^\lambda\big]^{\frac{1}{\lambda}},$$

$$d_H(\tilde{a}_H, \tilde{b}_H) = |a-b| + \tfrac{1}{m}\sum_{j=1}^{m} |\eta_{(j)} - \lambda_{(j)}|,$$

$$d_E(\tilde{a}_H, \tilde{b}_H) = \sqrt{|a-b|^2 + \tfrac{1}{m}\sum_{j=1}^{m} |\eta_{(j)} - \lambda_{(j)}|^2},$$

where $.{}_{(t)}$ is the tth largest value in its corresponding set.
If the real part of an arbitrary HFN is doubted and expressed by a HFE, then the HFN is extended to what named generalized hesitant fuzzy number (GHFN).

Definition 6: [7] Let $\mathbb{R}$ be the reference set. Then $\tilde{\tilde{E}}^H = \big\langle \{e_1, e_2, \ldots, e_m\}; \{\eta_1, \eta_2, \ldots, \eta_n\}\big\rangle$ is called a GHFN, where $g_i,\ i = 1, 2, \ldots, m$ are positive real quantities, and $\eta_j \in [0, 1],\ j = 1, 2, \ldots, n$ are membership/satisfaction degrees.

As it is seen, a GHFN similar to a HFN contain two parts: real values part, and hesitation degrees part. For any two GHFNs, if cardinalities of their real values parts are equal, and simultaneously the cardinalities of their hesitation degrees parts are also the same, they are called adjusted GHFNs (AGHFNs).

The mean value, score and variance of GHFNs can be obtained as follows.

Definition 7: [7] Let $\tilde{\tilde{E}}^H = \big\langle \{e_1, e_2, \ldots, e_m\}; \{\eta_1, \eta_2, \ldots, \eta_n\}\big\rangle$ be a GHFNs. Then, its mean value $\big(M(\tilde{\tilde{E}}^H)\big)$, its score $\big(Score(\tilde{\tilde{E}}^H)\big)$, and its variance $\big(Var(\tilde{\tilde{E}}^H)\big)$ can be obtained as follows:

$$M(\tilde{\tilde{E}}^H) = \big(\overline{e}, \overline{\eta}\big) = \left(\frac{\sum_{i=1}^{m} e_i}{m}, \frac{\sum_{i=1}^{n} \eta_i}{n}\right), \tag{1}$$

$$Score(\tilde{\tilde{E}}^H) = \overline{g} \times \overline{\eta} = \frac{\sum_{i=1}^{m} e_i}{m} \times \frac{\sum_{i=1}^{n} \eta_i}{n}, \tag{2}$$

$$Var(\tilde{\tilde{E}}^H) = \sqrt{\frac{\sum_{i=1}^{m} (e_i - \overline{e})^2}{m} + \sum_{i \neq j} (\eta_i - \eta_j)^2}. \tag{3}$$

Definition 8: [7] Let $w > 0$, $\tilde{\tilde{E}}^H = \big\langle \{e_1, e_2, \ldots, e_m\}; \{\eta_1, \eta_2, \ldots, \eta_n\}\big\rangle$ and $\tilde{\tilde{F}}^H = \big\langle \{f_1, f_2, \ldots, f_m\}; \{\lambda_1, \lambda_2, \ldots, \lambda_n\}\big\rangle$ be two AGHFNs. Then

$$(i) w\tilde{\tilde{E}}^H = \big\langle \{we_1, we_2, \ldots, we_m\}; \{\eta_1, \eta_2, \ldots, \eta_n\}\big\rangle; \tag{4}$$

$$(ii) \big(\tilde{\tilde{E}}^H\big)^w = \big\langle \{(e_1)^w, (e_2)^w, \ldots, (e_m)^w\}; \{\eta_1, \eta_2, \ldots, \eta_n\}\big\rangle; \tag{5}$$

$$(iii) \tilde{\tilde{E}}^H \oplus \tilde{\tilde{F}}^H = \big\langle \cup_i \{e_{(i)} + f_{(i)}\}; mh(E) \cup mh(F)\big\rangle; \tag{6}$$

$$(iv) \tilde{\tilde{E}}^H \otimes \tilde{\tilde{F}}^H = \big\langle \cup_i \{e_{(i)} f_{(i)}\}; \cup_i \min\{\eta_i, \lambda_i\}\big\rangle; \tag{7}$$

$$\tag{8}$$

where $\{e_{(1)}, e_{(2)}, \ldots, e_{(m)}\}$ is a reordering of $\{e_1, e_2, \ldots, e_m\}$ in which $e_{(1)} \leq e_{(2)} \leq \ldots \leq e_{(m)}$, and $\{f_{(1)}, f_{(2)}, \ldots, f_{(m)}\}$ is a permutation of $\{f_1, f_2, \ldots, f_m\}$ such that $f_{(1)} \leq f_{(2)} \leq \ldots \leq f_{(m)}$.

III. The proposed distance measures

In this Section, the new modified distance measures will be defined.

Definition 9: Consider two arbitrary given adjusted HFNs $\tilde{a}_H = \langle a; \{\eta_1, \eta_2, \ldots, \eta_m\}\rangle$, and $\tilde{b}_H = \langle b; \{\lambda_1, \lambda_2, \ldots, \lambda_m\}\rangle$. Then, the new modification of Hamming hesitant distance $d_{MH}(\tilde{a}_H, \tilde{b}_H)$, Euclidean hesitant distance $d_{ME}(\tilde{a}_H, \tilde{b}_H)$, and Generalized hesitant distance $d_{MG}(\tilde{a}_H, \tilde{b}_H)$ will be defined as follows:

$$d_{MH}(\tilde{a}_H, \tilde{b}_H) = |a-b| + \tfrac{1}{m^2}\sum_{\eta_j, \lambda_j} \big|\tfrac{\eta_j - \lambda_j}{\eta_j + \lambda_j}\big|,$$

$$d_{ME}(\tilde{a}_H, \tilde{b}_H) = \sqrt{|a-b|^2 + \tfrac{1}{m^2}\sum_{\eta_j, \lambda_j} \big|\tfrac{\eta_j - \lambda_j}{\eta_j + \lambda_j}\big|^2}, \tag{9}$$

$$d_{MG}(\tilde{a}_H, \tilde{b}_H) = \big[|a-b|^\lambda + \tfrac{1}{m^2}\sum_{\eta_j, \lambda_j} \big|\tfrac{\eta_j - \lambda_j}{\eta_j + \lambda_j}\big|^\lambda\big]^{\frac{1}{\lambda}}.$$

where $\lambda > 0$.
In some situations, the two real and membership parts of given HFNs are not equal importants.

Definition 10: For any two arbitrary given adjusted HFNs $\tilde{a}_H = \langle a; \{\eta_1, \eta_2, \ldots, \eta_m\}\rangle$, and $\tilde{b}_H = \langle b; \{\lambda_1, \lambda_2, \ldots, \lambda_m\}\rangle$, where their real and membership parts are weighted by $0 \leq w \leq 1$ and $0 \leq 1 - w \leq 1$, respectively. Then, the weighted distance measures can be defined as follows.

$$d_{WMH}(\tilde{a}_H, \tilde{b}_H) = w|a-b| + (1-w)\tfrac{1}{m^2}\sum_{\eta_j, \lambda_j} \big|\tfrac{\eta_{(j)} - \lambda_{(j)}}{\eta_j + \lambda_j}\big|,$$

$$d_{WME}(\tilde{a}_H, \tilde{b}_H) = \sqrt{w|a-b|^2 + (1-w)\tfrac{1}{m^2}\sum_{\eta_j, \lambda_j} \big|\tfrac{\eta_j - \lambda_j}{\eta_j + \lambda_j}\big|^2},$$

$$d_{WMG}(\tilde{a}_H, \tilde{b}_H) = \big[w|a-b|^\lambda + (1-w)\tfrac{1}{m^2}\sum_{\eta_j, \lambda_j} \big|\tfrac{\eta_j - \lambda_j}{\eta_j + \lambda_j}\big|^\lambda\big]^{\frac{1}{\lambda}}. \tag{10}$$

IV. Numerical example

In some real practical problems, we have reference points as patterns, and based on minimum difference from reference points, any other given points should be checked to which subject it belongs.

Let $\tilde{a}_H^1 = \langle 124; \{.9, .8, .7\}\rangle, \tilde{a}_H^2 = \langle 124; \{.2, .6, .7\}\rangle$ and $\tilde{a}_H^3 = \langle 124; \{.7, .65, .6\}\rangle$ be three given HFNs. As it can be seen, these HFNs are different only with their membership parts. Based on (5), we have $d_H(\tilde{a}_H^1, \tilde{a}_H^3) = d_H(\tilde{a}_H^2, \tilde{a}_H^3) = 0.15$. It is showed that HFN $\tilde{a}_H^3$ belongs to both patterns $\tilde{a}_H^1$ and $\tilde{a}_H^2$, while unlike $\tilde{a}_H^2$, HFN $\tilde{a}_H^1$ has small divegance in its membership part. Then we can not determine that $\tilde{a}_H^3$ belongs to $\tilde{a}_H^1$ or $\tilde{a}_H^2$ patterns.

Using the proposed method in (10), we will have $d_H(\tilde{a}_H^1, \tilde{a}_H^3) = 0.1017, \quad d_H(\tilde{a}_H^2, \tilde{a}_H^3) = 0.18$. It means, HFN $\tilde{a}_H^3$, as its minimum distance from HFN $\tilde{a}_H^1$, belongs to pattern defined by HFN $\tilde{a}_H^1$.

V. Conclusion

In this paper new modified distance measures of Hamming distance, Euclidean distance, and generalized distance measures have been proposed. In the future, multi-attribute decision making problems and some other practical problems such as medical diagnostic problems, transportation and allocation problems, social programming, etc, will be tried to be solved using these new tools for modeling uncertainty.

References

[1] K. T. Atanassov," Intuitionistic fuzzy sets," in: V. Sgurev, Ed., VII ITKR'S Session, Sofia, Jone 1983.

[2] R. Chinram, A. Hussain, T. Mahmood and M. I. Ali, "EDAS method for multi-criteria group decision making based on intuitionistic fuzzy rough aggregation operators," in IEEE Access, vol. 9, pp. 10199–10216, 2021.

[3] T. Denoeux, "Dempster-Shafer theory: Introduction, connections with rough sets and application to clustering," invited talk, 9th International Conference on Rough Sets and Knowledge Technology (RSKT 2014), Shanghai, China, October 24-26, 2014.

[4] H. Garg, A. Keikha, H. M. Nehi, Multiple attribute decision making problem using TOPSIS and Choquet integral with hesitant fuzzy number information, Math. Probl. Eng., 2020, https://doi.org/10.1155/2020/9874951.

[5] N. N. Karnik and J. M. Mendel, "Centroid of a type-2 fuzzy set," J. Inf. Sci., vol. 132, pp. 195–220, 2001.

[6] A. Keikha, "Archimedean t-norm and t-conorm-based aggregation operators of HFNs, with the approach of improving education," Int. J. Fuzzy Syst., 2021, https://doi.org/10.1007/s40815-021-01137-3.

[7] A. Keikha, "Generalized hesitant fuzzy numbers: introducing, arithmetic operations, aggregation operators, and an application," Int. J. Intell. Syst., vol. 36, 12, pp. 7709–7730, 2021.

[8] A. Keikha, "Introducing a new type of HFSs and their application in solving MADM," J. Intell. Fuzzy Syst., vol. 40, pp. 9333–9344 , 2021.

[9] EP. Klement, R. Mesiar, Logical, Algebraic, Analytic, and Probabilistic Aspects of Triangular Norms. Elsevier, New York, 2005.

[10] G. J. Klir, Uncertainty and Information : Foundations of Generalized Information Theory. John Wiley & Sons, Inc., 2006.

[11] G. Klir, B. Yuan, Fuzzy Sets and Fuzzy Logic: Theory and Applications. Prentice Hall, Upper Saddle River, 1995.

[12] H. C. Liao and Z. S. Xu, Hesitant Fuzzy Decision Making Methodologies and Applications. Springer, Singapore, 2017.

[13] T. Mahmood, TW. Ali, Z. Ali, R. Chinram, "Power aggregation operators and similarity measures based on improved intuitionistic hesitant fuzzy sets and their applications to multiple attribute decision making," Comput Model Eng Sci, vol. 126, 3, pp. 1165–1187, 2021.

[14] B. Singha, M. Sen, N. Sinha, "Modified distance measure on hesitant fuzzy sets and its application in multi-criteria decision making problem", OPSEARCH, Springer, Operational Research Society of India, vol. 57, 2, pp. 584–602, June, 2020.

[15] V. Torra, "Hesitant fuzzy sets," Int. J. Intell. Syst., vol. 25, 1, pp. 529–539, 2010.

[16] L. Wang, H. Garg, "Algorithm for multiple attribute decision making with interactive Archimedean norm operations under Pythagorean fuzzy uncertainty," Int. J. Comput. Intell., vol. 14, 1, pp. 503–527, 2020.

[17] W. Weaver, "Science and complexity," Am. Sci., vol. 36, pp. 536–544, 1948.

[18] L. A. Zadeh, "Fuzzy sets," Inf. and Cont., vol. 8, pp. 338–353, 1965.

[19] Z. Zhang, "Several new hesitant fuzzy aggregation operators and their application to multi-criteria decision making," Proceedings of the National Academy of Sciences, India - Section A 86(3), 2016.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

True-False triangular norm and conorms and It's Application

1st M. Mohseni Takallo, 2nd Y.B. Jun and 3rd R.A. Borzooei

1st Department of Mathematics, *Shahid Beheshti University*, Tehran, Iran, mohammad.mohseni1122@gmail.com
2nd Department of Mathematics Education, *Gyeongsang National University*, Jinju 52828, Korea, skywine@gmail.com
3rd Department of Mathematics, *Shahid Beheshti University*, Tehran, Iran, borzooei@sbu.ac.ir

***Abstract*— In this article, we introduced triangular norm $\mathfrak{T}$ (briefly t-norm) and triangular co-norm $\mathfrak{S}$ (briefly s-norm) of combination a fuzzy set and an interval-valued fuzzy set and proposed $\mathfrak{T}$-representable and $\mathfrak{S}$-representable of them. By using $\mathfrak{T}$-representable and $\mathfrak{S}$-representable, we suggested $(\mathfrak{T},\mathfrak{S})$-representable of true-false sets (TFS). We defined the TF implication, the residuated TF implication and considered property between residuated TF implication with Lukasiewicz t-norm and Hamming distance, Hausdorff distance, Chebyshev distance and Moore distance. Also, we defined the degree of similarity between two TFS and considered property between the degree of similarity with some type of distance. Later on, we extended the TOPSIS method based on TFS. Finally, we applied proposed the TOPSIS method solve the decision-making problem.**

***Index Terms*—$\mathfrak{T}$-representable, $\mathfrak{S}$-representable, $(\mathfrak{T},\mathfrak{S})$-representable of TFS, TF implication, Residuated TF implication, True-False Topsis.**

I. Introduction

In 1965, Zadeh [27] introduced the fuzzy set theory (FS) that has shown meaningful applications in many fields of study. The idea of fuzzy set is welcome because it handles uncertainty and vagueness which Cantorian set could not address. In fuzzy set theory, the membership of an element to a fuzzy set is a single value between zero and one. As a generalization of fuzzy set, Zadeh [28–30] made an extension of the concept of a fuzzy set by an interval-valued fuzzy set (IVFS), i.e., a fuzzy set with an intervalvalued membership function. In 2018, Mohseni et al.[17] ntroduced the notion of MBJ-neutrosophic sets which is another generalization of neutrosophicset. They use the interval valued fuzzy set as the indeterminate membership function because they claim interval-valued fuzzy set can be better-shown hesitancy as an indeterminate membership function and also, interval valued fuzzy set is a generalization of a fuzzy set. Later on, Borzooei et al. [4] introduced the concept of true-false sets (TFS) and claimed that true-false set is generalization of fuzzy sets, intuitionistic fuzzy set, vague sets, interval-valued fuzzy sets, cubic sets, neutrosophic sets, MBJ-neutrosophic set and etc. They applied TFS to groups and BCK/BCI-algebras [18–20]. The history of triangular norms started with the paper "Statistical metrics" [16]. The main idea of Karl Menger was to construct metric spaces where probability distributions rather than numbers are used in order to describe the distance between two elements of the space in question. Triangular norms (t-norms for short) naturally came into the picture in the course of the generalization of the classical triangle inequality to this more general setting. The original set of axioms for t-norms was considerably weaker, including among others also the functions which are known today as triangular co-norms. Consequently, the first field where t-norms played a major role was the theory of probabilistic metric spaces (as statistical metric spaces were called after 1964). Schweizer et al. [23] provided the axioms oft-norms, as they are used today, and a redefinition of statistical metric spaces given in [24] led to a rapid development of the field. Many results concerning t-norms were obtained in the course of this development, most of which are summarized in the monograph [22]. Triangular norms can be defined on any bounded poset [7]. Deschrijver et al. [8–11] discussed some classes of interval-valued t-norms and intuitionistic t-norms.

In [2, 3], Fuzzy systems make extensive use of if-then rules, also known as Horn clauses. Classically, any if-then rule is modelled by a classical implication. Distance measures have become important due to the significant applications in diverse fields like remote sensing, data mining, pattern recognition [**?**] and multivariate data analysis. Several distance measures for precise numbers are well established in the literature. But a logical problem arises as the distance is computed in an imprecise framework due to the existence of vagueness.

II. Preliminaries

A *fuzzy set* in a set X is defined to be a function $\lambda : X \to I$ where $I = [0, 1]$. Denote by I^X the collection of all fuzzy sets

in a set X [27]. Define a relation $\leq$ on I^X as follows:

$$(\forall \lambda, \mu \in I^X)\,(\lambda \leq \mu \iff (\forall x \in X)(\lambda(x) \leq \mu(x))).$$

The join ($\vee$) and meet ($\wedge$) of λ and μ are defined by

$$(\lambda \vee \mu)(x) = \max\{\lambda(x), \mu(x)\},$$
$$(\lambda \wedge \mu)(x) = \min\{\lambda(x), \mu(x)\},$$

respectively, for all $x \in X$. The complement of λ, denoted by λ^c, is defined by

$$(\forall x \in X)\,(\lambda^c(x) = 1 - \lambda(x)).$$

For a family $\{\lambda_i \mid i \in \Lambda\}$ of fuzzy sets in X, we define the join ($\vee$) and meet ($\wedge$) operations as follows:

$$\Big(\bigvee_{i\in\Lambda} \lambda_i\Big)(x) = \sup\{\lambda_i(x) \mid i \in \Lambda\},$$
$$\Big(\bigwedge_{i\in\Lambda} \lambda_i\Big)(x) = \inf\{\lambda_i(x) \mid i \in \Lambda\},$$

respectively, for all $x \in X$.

By an *interval number* we mean a closed subinterval $\tilde{a} := [a^-, a^+]$ of $[0,1]$, where $0 \leq a^- \leq a^+ \leq 1$. Denote by $int([0,1])$ the set of all interval numbers. Let us define what is known as *refined minimum* (briefly, rmin) and *refined maximum* (briefly, rmax) of two elements in $int([0,1])$. We also define the symbols "$\succcurlyeq$", "$\preccurlyeq$", "$=$" in case of two elements in $int([0,1])$. Consider two interval numbers $\tilde{a}_1 := \left[a_1^-, a_1^+\right]$ and $\tilde{a}_2 := \left[a_2^-, a_2^+\right]$. Then

$$\operatorname{rmin}\{\tilde{a}_1, \tilde{a}_2\} = \left[\min\left\{a_1^-, a_2^-\right\}, \min\left\{a_1^+, a_2^+\right\}\right], \quad (1)$$
$$\operatorname{rmax}\{\tilde{a}_1, \tilde{a}_2\} = \left[\max\left\{a_1^-, a_2^-\right\}, \max\left\{a_1^+, a_2^+\right\}\right], \quad (2)$$
$$\tilde{a}_1 \succcurlyeq \tilde{a}_2 \text{ if and only if } a_1^- \geq a_2^- \text{ and } a_1^+ \geq a_2^+, \quad (3)$$
$$\tilde{a}_1 \preccurlyeq \tilde{a}_2 \text{ if and only if } a_1^- \leq a_2^- \text{ and } a_1^+ \leq a_2^+, \quad (4)$$

and similarly we may have $\tilde{a}_1 \preccurlyeq \tilde{a}_2$ and $\tilde{a}_1 = \tilde{a}_2$. To say $\tilde{a}_1 \succ \tilde{a}_2$ (resp. $\tilde{a}_1 \prec \tilde{a}_2$) we mean $\tilde{a}_1 \succcurlyeq \tilde{a}_2$ and $\tilde{a}_1 \neq \tilde{a}_2$ (resp. $\tilde{a}_1 \preccurlyeq \tilde{a}_2$ and $\tilde{a}_1 \neq \tilde{a}_2$). Let $\tilde{a}_i \in int([0,1])$ where $i \in \Lambda$. We define

$$\operatorname*{rinf}_{i\in\Lambda} \tilde{a}_i = \left[\inf_{i\in\Lambda} a_i^-, \inf_{i\in\Lambda} a_i^+\right] \quad \text{and} \quad \operatorname*{rsup}_{i\in\Lambda} \tilde{a}_i = \left[\sup_{i\in\Lambda} a_i^-, \sup_{i\in\Lambda} a_i^+\right].$$

For any $\tilde{a} \in int([0,1])$, its *complement*, denoted by $\tilde{a}^c$, is defined to be the interval number

$$\tilde{a}^c = [1 - a^+, 1 - a^-].$$

Let X be a nonempty set. A function $A : X \to [I]$ is called an *interval-valued fuzzy set* (briefly, an *IVF set*) in X [28–30]. Let $[I]^X$ stand for the set of all IVF sets in X. For every $A \in [I]^X$ and $x \in X$, $A(x) = [A^-(x), A^+(x)]$ is called the *degree* of membership of an element x to A, where $A^- : X \to I$ and $A^+ : X \to I$ are fuzzy sets in X which are called a *lower fuzzy set* and an *upper fuzzy set* in X, respectively. For simplicity, we denote $A = [A^-, A^+]$. We define $\mathbf{1}^* = [1,1]$ is the greatest interval-valued fuzzy set on X, and $\mathbf{0}^* = [0,0]$ is the smallest interval-valued fuzzy set on X and for every $A, B \in [I]^X$,

$$A \subseteq B \Leftrightarrow A(x) \preceq B(x) \text{ for all } x \in X,$$

and

$$A = B \Leftrightarrow A(x) = B(x) \text{ for all } x \in X.$$

The complement A^c of $A \in [I]^X$ is defined as follows: $A^c(x) = A(x)^c$ for all $x \in X$, that is,

$$A^c(x) = [1 - A^+(x), 1 - A^-(x)] \text{ for all } x \in X.$$

For a family $\{A_i \mid i \in \Lambda\}$ of IVF sets in X where Λ is an index set, the *union* $G = \bigcup_{i\in\Lambda} A_i$ and the *intersection* $F = \bigcap_{i\in\Lambda} A_i$ are defined as follows:

$$G(x) = \left(\bigcup_{i\in\Lambda} A_i\right)(x) = \operatorname*{rsup}_{i\in\Lambda} A_i(x)$$

and

$$F(x) = \left(\bigcap_{i\in\Lambda} A_i\right)(x) = \operatorname*{rinf}_{i\in\Lambda} A_i(x)$$

for all $x \in X$, respectively. For a point $p \in X$ and for $\tilde{a} = [a^-, a^+] \in [I]$ with $a^+ > 0$, the IVF set which takes the value $\tilde{a}$ at p and $\mathbf{0}$ elsewhere in X is called an *interval-valued fuzzy point* (briefly, an *IVF point*) and is denoted by $\tilde{a}_p$. The set of all IVF points in X is denoted by $IVFP(X)$. For any $\tilde{a} \in [I]$ and $x \in X$, the IVF point $\tilde{a}_x$ is said to *belong* to an IVF set A in X, denoted by $\tilde{a}_x \tilde{\in} A$, if $A(x) \succeq \tilde{a}$. It can be easily shown that $A = \cup\{\tilde{a}_x \mid \tilde{a}_x \tilde{\in} A\}$.

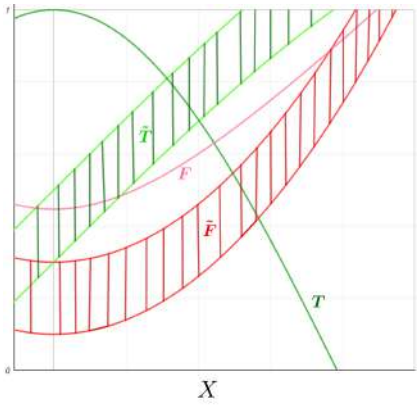
Definition 1. *Let X be a non-empty set. By a True-False sets (briefly, TFS) in X, we mean a structure of the form:*

$$\mathcal{A} := \{\langle x; t_A(x), \tilde{T}_A(x), f_A(x), \tilde{F}_A(x)\rangle \mid x \in X\}$$

where t and $\tilde{T}$ are fuzzy sets and interval-valued fuzzy set respectively in X, which are called a truth membership function. Also, f and $\tilde{F}$ are fuzzy sets and interval-valued fuzzy set respectively in X, which are called false membership function. In other words,

$$\tilde{T}_A, \tilde{F}_A : X \to \mathcal{I}[0,1]$$

where $\mathcal{I}[0,1]$ is the set of all sub-intervals of $[0,1]$. We define $TF(X)$ as all TF-set of X.



Note 1. *An TF value is denoted by $(t_A, \tilde{T}_A, f_A, \tilde{F}_A)$ for convenience.*

Definition 2. *[14, 26] A triangular norm (t-norm for short)(resp. s-norm) is defined to be a function* $T : [0,1] \times [0,1] \longrightarrow [0,1]$*(resp.* $S : [0,1] \times [0,1] \longrightarrow [0,1]$*), where for every* $u, v, w, x \in [0,1]$*, it satisfies the following properties (axioms):*
(T1) $T(u,1) = u$ *(resp.* $S(u,0) = u$*)*
(T2) $T(u,v) = T(v,u)$ *(resp.* $S(u,v) = S(v,u)$*)*
(T3) $T(T(u,v),w) = T(u,T(v,w))$ *(resp.* $S(S(u,v),w) = S(u,S(v,w))$*)*
(T4) If $u \leqslant x$ *and* $v \leqslant w$*, then* $T(u,v) \leqslant T(x,w)$ *(resp.* $S(u,v) \leqslant S(x,w)$*)*

Notation 1. *Let* T *be a t-norm. Since* T *has assosiative property, for every* $u \in [0,1]$*, we denoted:*

$$\begin{aligned} T(u,u) &= u_T^{(2)} \\ T(u,T(u,u)) &= T(u,u,u) = u_T^{(3)} \\ &\vdots \\ T(u_T^{(n-1)},u) &= T(u,u,...,u) = u_T^n \end{aligned}$$

Definition 3. *[14] A t-norm* T *is called left-continuous, if for any* $(u_0,v_0) \in [0,1] \times [0,1]$ *and for each* $\varepsilon > 0$*, there is a* $\delta > 0$ *such that* $T(u,v) > T(u_0,v_0) - \varepsilon$ *whenever* $(u,v) \in (u_0 - \delta, u_0] \times (v_0 - \delta, v_0]$.

Definition 4. *[2] A mapping* $\Rightarrow$: $[0,1] \times [0,1] \longrightarrow [0,1]$ *is called a fuzzy implication if it is decreasing on the first variable and increasing on the second one and it satisfies*

$$0 \Rightarrow 0 = 1, 1 \Rightarrow 1 = 1 \text{ and } 1 \Rightarrow 0 = 0.$$

Also, if the fuzzy implication $\Rightarrow$ *satisfies to*
(i) *If* $x \leq y$*, then* $x \Rightarrow z \geq y \Rightarrow z$*, for all* $z \in [0,1]$,
(ii) *If* $x \leq y$*, then* $z \Rightarrow x \leq z \Rightarrow y$*, for all* $z \in [0,1]$,
then $\Rightarrow$ *is called the strong fuzzy implication.*

Definition 5. *Let* T *be a t-norm, then mapping* $\Rightarrow_T$: $[0,1] \times [0,1] \longrightarrow [0,1]$ *is called the residuated implication and defined as follows:*

$$x \Rightarrow_T y = \bigvee \{z \in [0,1] | T(x,z) \leq y\}$$

Note 2. *Easy to see, the residuated implication is fuzzy implication.*

Definition 6. *[10] A t-norm (resp. s-norm) on* $[I]$ *is defined to be a function* $\mathcal{T} : [I] \times [I] \longrightarrow [I]$ *(resp.* $\mathcal{S} : [I] \times [I] \longrightarrow [I]$*), where for every* $\tilde{u}, \tilde{v}, \tilde{w}, \tilde{x} \in [I]$*, it satisfies the following properties (axioms):*
(T1) $\mathcal{T}(\tilde{u}, \mathbf{1}^*) = \tilde{u}$ *(resp.* $\mathcal{S}(\tilde{u}, \mathbf{0}^*) = \tilde{u}$*)*
(T2) $\mathcal{T}(\tilde{u}, \tilde{v}) = \mathcal{T}(\tilde{v}, \tilde{u})$ *(resp.* $\mathcal{S}(\tilde{u}, \tilde{v}) = \mathcal{S}(\tilde{v}, \tilde{u})$*)*
(T3) $\mathcal{T}(\mathcal{T}(\tilde{u}, \tilde{v}), \tilde{w}) = \mathcal{T}(\tilde{u}, \mathcal{T}(\tilde{v}, \tilde{w}))$ *(resp.* $\mathcal{S}(\mathcal{S}(\tilde{u}, \tilde{v}), \tilde{w}) = \mathcal{S}(\tilde{u}, \mathcal{S}(\tilde{v}, \tilde{w}))$*)*
(T4) If $\tilde{u} \preceq \tilde{x}$ *and* $\tilde{v} \preceq \tilde{w}$*, then* $\mathcal{T}(\tilde{u}, \tilde{v}) \preceq \mathcal{T}(\tilde{x}, \tilde{w})$ *(resp.* $\mathcal{S}(\tilde{u}, \tilde{v}) \preceq \mathcal{S}(\tilde{x}, \tilde{w})$*)*

Definition 7. *[10, 11] A t-norm* $\mathcal{T}_{T_1,T_2}$ *on* $[I]$ *is called t-representable if there exist t-norms* T_1 *and* T_2 *on the unit interval*$[0,1]$ *such that* $T_1(x,y)T_2(x,y)$*, for all* $x, y \in [0,1]$*, and for all* $[a,b], [c,d] in [I]$, $\mathcal{T}_{T_1,T_2} : [I] \times [I] \longrightarrow [I]$ *is defined as*

$$\mathcal{T}_{T_1,T_2}([a,b],[c,d]) = [T_1(a,c), T_2(b,d)].$$

Note 3. *If* $T_1 = T_2 = T$*, then* $\mathcal{T}_{T_1,T_2}([a,b],[c,d]) = T([a,b],[c,d]) = [T(a,c), T(b,d)]$ *is the associated t-norm on* $[I]$.

Definition 8. *The interval-valued t-representable t-norm* $\mathcal{T}_{T_1,T_2}$ *is called a left-continuous if* T_1 *and* T_2 *are left-continuous respectively.*

Definition 9. *[?] A mapping* $\Rightarrow_I$: $[I] \times [I] \longrightarrow [I]$ *is called an interval-valued fuzzy implication if it is non-increasing on the first component and non-decreasing on the second component and satisfies* $\mathbf{0}^* \Rightarrow_I \mathbf{0}^* = \mathbf{1}^*, \mathbf{0}^* \Rightarrow_I \mathbf{1}^* = \mathbf{1}^*, \mathbf{1}^* \Rightarrow_I \mathbf{1}^* = \mathbf{1}^*, \mathbf{1}^* \Rightarrow_I \mathbf{0}^* = \mathbf{0}^*$.

Definition 10. *[11] Let* $\mathcal{T}_{T_1,T_2}$ *be a t-representable t-norm, then mapping* $\Rightarrow_{\mathcal{T}}$: $[I] \times [I] \longrightarrow [I]$ *is called the residuated implication and defined as follows:*

$$\tilde{x} \Rightarrow_\tau \tilde{y} = \bigvee \{\tilde{z} \in [I] | \mathcal{T}_{T_1,T_2}(\tilde{x}, \tilde{z}) \preceq \tilde{y}\}$$

III. True-False triangular norm and conorms

In this section, we describe some concepts and an important theorem that will be required for our following work.

Definition 11. *A t-norm (resp. s-norm) is defined to be a function* $\mathfrak{T} : (I,[I]) \times (I,[I]) \longrightarrow (I,[I])$ *(resp.* $\mathfrak{S} : (I,[I]) \times (I,[I]) \longrightarrow (I,[I])$ *), where for every* $(u,\tilde{u}), (v,\tilde{v}), (w,\tilde{w}), (x,\tilde{x}) \in (I,[I])$*, it satisfies the following properties (axioms):*
(T1) $\mathfrak{T}((u,\tilde{u}),(1,\mathbf{1}^*)) = (u,\tilde{u})$ *(resp.* $\mathfrak{S}((u,\tilde{u}),(0,\mathbf{0}^*)) = (u,\tilde{u})$*)*
(T2) $\mathfrak{T}((u,\tilde{u}),(v,\tilde{v})) = \mathfrak{T}((v,\tilde{v}),(u,\tilde{u}))$ *(resp.* $\mathfrak{S}((u,\tilde{u}),(v,\tilde{v})) = \mathfrak{S}((v,\tilde{v}),(u,\tilde{u}))$*)*
(T3) $\mathfrak{T}(\mathfrak{T}((u,\tilde{u}),(v,\tilde{v})),(w,\tilde{w}))$=$\mathfrak{T}((u,\tilde{u}),\mathfrak{T}((v,\tilde{v}),(w,\tilde{w})))$
(resp. $\mathfrak{S}(\mathfrak{S}((u,\tilde{u}),(v,\tilde{v})),(w,\tilde{w}))$=$\mathfrak{S}((u,\tilde{u}),\mathfrak{S}((v,\tilde{v}),(w,\tilde{w}))))$
(T4) If $(u,\tilde{u}) \ll (x,\tilde{x})$ *and* $(v,\tilde{v}) \ll (w,\tilde{w})$*, then* $\mathfrak{T}((u,\tilde{u}),(v,\tilde{v})) \ll \mathfrak{T}((x,\tilde{x}),(w,\tilde{w}))$
(resp. $\mathfrak{S}((u,\tilde{u}),(v,\tilde{v})) \ll \mathfrak{S}((x,\tilde{x}),(w,\tilde{w}))$*)*

Definition 12. *A t-norm* $\mathfrak{T}_{(T_1,\mathcal{T}_{T_2,T_3})}$ *on* $(I,[I])$ *is called* $\mathfrak{T}$*-representable if there exist t-norms* T_1 *and* $\mathcal{T}_{T_2,T_3}$ *on the unit interval* I *and close sub-interval* $[I]$*, respectively, such that* $T_2(x,y) \leq T_3(x,y)$*, for all* $x, y \in [0,1]$*, and for all* $\tilde{x} = [a,b], \tilde{y} = [c,d] in [I], \mathfrak{T}_{(T_1,\mathcal{T}_{T_2,T_3})} : (I,[I]) \times (I,[I]) \longrightarrow (I,[I])$ *is defined as*

$$\begin{aligned} \mathfrak{T}_{(T_1,\mathcal{T}_{T_2,T_3})}((x,\tilde{x}),(y,\tilde{y})) &= \mathfrak{T}_{(T_1,\mathcal{T}_{T_2,T_3})}((x,[a,b]),(y,[c,d])) \\ &= (T_1(x,y),[T_2(a,c),T_3(b,d)]). \end{aligned}$$

Definition 13. *A s-norm* $\mathfrak{S}_{(S_1,\mathcal{S}_{S_2,S_3})}$ *on* $(I,[I])$ *is called* $\mathfrak{S}$*-representable if there exist s-norms* S_1 *and* $\mathcal{S}_{S_2,S_3}$ *on the unit interval* I *and close sub-interval* $[I]$*, respectively, such that* $S_2(x,y) \leq S_3(x,y)$*, for all* $x, y \in [0,1]$*, and for all* $\tilde{x} = [a,b], \tilde{y} = [c,d] \in [I], \mathfrak{S}_{(S_1,\mathcal{S}_{S_2,S_3})} : (I,[I]) \times (I,[I]) \longrightarrow (I,[I])$ *is defined as*

$$\begin{aligned} \mathfrak{S}_{(S_1,\mathcal{T}_{S_2,S_3})}((x,\tilde{x}),(y,\tilde{y})) &= \mathfrak{S}_{(S_1,\mathcal{T}_{S_2,S_3})}((x,[a,b]),(y,[c,d])) \\ &= (S_1(x,y),[S_2(a,c),S_3(b,d)]). \end{aligned}$$

Definition 14. *A t-norm* $\mathfrak{TS} := (\mathfrak{T}_{(T_1,\mathcal{T}_{T_2,T_3})}, \mathfrak{S}_{(S_1,\mathcal{S}_{S_2,S_3})})$ *of* TFS *is called* $(\mathfrak{T},\mathfrak{S})$*-representable , for all* $x, y, z, w \in [0,1]$,

and for all $\tilde{x} = [a, b], \tilde{y} = [c, d], \tilde{z} = [e, f], \tilde{w} = [g, h] \in [I]$, *is defined as*

$$\mathfrak{TS} : ((I,[I]),(I,[I]))\times((I,[I]),(I,[I])) \longrightarrow ((I,[I]),(I,[I]))$$

$$\mathfrak{TS}(((x,\tilde{x},y,\tilde{y})),((z,\tilde{z},w,\tilde{w})))$$
$$= ((T_1(x,z),[T_2(a,e),T_3(b,f)]),(S_1(y,w),[S_2(c,g),S_3(d,h)])).$$

Definition 15. *A mapping* $\Rightarrow_{TF}$: $((I,[I]),(I,[I])) \times ((I,[I]),(I,[I])) \longrightarrow ((I,[I]),(I,[I]))$ *is called an* TF *implication if it satisfies in following properties:*

$$\begin{aligned} 0^\star \Rightarrow_{TF} 0^\star &= 1^\star, 0^\star \Rightarrow_{TF} 1^\star \\ &= 1^\star, 1^\star \Rightarrow_{TF} 1^\star \\ &= 1^\star, 1^\star \Rightarrow_{TF} 0^\star \\ &= 0^\star. \end{aligned}$$

Where $1^\star = (1,[1,1],0,[0,0])$ *and* $0^\star = (0,[0,0],1,[1,1])$ *are the greatest and the smallest* TFS *on* X, *respectively.*

Definition 16. *Let* $\mathfrak{TS}$ *be a* $(\mathfrak{T},\mathfrak{S})$*-representable t-norm, then mapping* $\Rightarrow_{TF}$: $((I,[I]),(I,[I]))\times((I,[I]),(I,[I])) \longrightarrow ((I,[I]),(I,[I]))$ *is called the residuated* TF *implication (briefly* $\mathcal{R}_{TF}$ *implication)and defined as follows:*

$$\begin{aligned} \mathcal{R}_{\mathfrak{TS}} :&= (x,\tilde{x},y,\tilde{y}) \Rightarrow_{TF} (z,\tilde{z},w,\tilde{w}) \\ &= \bigvee\{(a,\tilde{a},b,\tilde{b}) | \mathfrak{TS}((x,\tilde{x},y,\tilde{y}),(a,\tilde{a},b,\tilde{b})) \\ &\sqsubseteq (z,\tilde{z},w,\tilde{w})\} \end{aligned}$$

Definition 17. *Let* X *be non-empty set. For any* $\mathcal{A},\mathcal{B} \in TF(X)$ *we define*

(i) *The Hamming distance between* $\mathcal{A}$ *and* $\mathcal{B}$ *as follows:*

$$\begin{aligned} d(\mathcal{A},\mathcal{B}) = \sum_{i=1}^{n} &(|t_A(x_i) - t_B(x_i)| \\ &+ |\tilde{T}_A(x_i) - \tilde{T}_B(x_i)| + |f_A(x_i) - f_B(x_i)| \\ &+ |\tilde{F}_A(x_i) - \tilde{F}_B(x_i)|) \end{aligned} \tag{5}$$

(ii) *The Hausdorff distance between* $\mathcal{A}$ *and* $\mathcal{B}$ *as follows:*

$$\begin{aligned} d(\mathcal{A},\mathcal{B}) = \frac{1}{n}\sum_{i=1}^{n} \big\{ &\max\{|t_A(x_i) \\ &- t_B(x_i)|, |\tilde{T}_A(x_i) - \tilde{T}_B(x_i)|, |f_A(x_i) \\ &- f_B(x_i)|, |\tilde{F}_A(x_i) - \tilde{F}_B(x_i)|\}\big\} \end{aligned} \tag{6}$$

(iii) *The Chebyshev distance between* $\mathcal{A}$ *and* $\mathcal{B}$ *as follows:*

$$\begin{aligned} d(\mathcal{A},\mathcal{B}) = &\sup_i |t_A(x_i) - t_B(x_i)| \\ &+ \sup_i |\tilde{T}_A(x_i) - \tilde{T}_B(x_i)| \\ &+ \sup_i |f_A(x_i) - f_B(x_i)| \\ &+ \sup_i |\tilde{F}_A(x_i) - \tilde{F}_B(x_i)| \end{aligned} \tag{7}$$

(iv) *The Moore distance between* $\mathcal{A}$ *and* $\mathcal{B}$ *as follows:*

$$\begin{aligned} d(\mathcal{A},\mathcal{B}) = &\max_i |t_A(x_i) - t_B(x_i)| \\ &+ \max\{\max_i |T_A(x_i)^- - T_B(x_i)^-|, \max_i |T_A(x_i) \\ &+ -T_B(x_i)^+|\} + \min_i |f_A(x_i) - f_B(x_i)| \\ &+ \min\{\min_i |F_A(x_i)^- - F_B(x_i)^-|, \\ &\min_i |F_A(x_i)^+ - F_B(x_i)^+|\} \end{aligned} \tag{8}$$

The operation $\zeta_{\Rightarrow_{TF}} = \mathcal{R}_{\mathfrak{TS}}(\tilde{xy},\tilde{zw}) \overset{\propto}{\wedge} \mathcal{R}^{-1}_{\mathfrak{TS}}(\tilde{xy},\tilde{zw})$ is called $\overset{\propto}{\wedge} - \mathcal{R}_{TF}$ implication and defined by

$$\begin{aligned} \tilde{xy} \overset{\propto}{\wedge} \tilde{zw} =&(x,\tilde{x},y,\tilde{y}) \overset{\propto}{\wedge} (z,\tilde{z},w,\tilde{w}) \\ =&(x\wedge z, \tilde{x}\tilde{\wedge}\tilde{z}, y\vee w, \tilde{y}\tilde{\vee}\tilde{w}) \end{aligned}$$

and where $\tilde{xy}$ and $\tilde{zw}$ are $(x,\tilde{x},y,\tilde{y})$ and $(z,\tilde{z},w,\tilde{w})$, respectively.

Definition 18. *Let* $\zeta_{\Rightarrow_{TF}}$ *be a* $\overset{\propto}{\wedge} - \mathcal{R}_{TF}$ *implication. The degree of similarity or degree of equality from any* $\mathcal{A},\mathcal{B} \in TF(X)$ *denoted by* $\mathcal{SI}_{TF}(\mathcal{A}(x),\mathcal{B}(x))$ *and defined as following:*

$$\begin{aligned} \mathcal{SI}_{TF}(\mathcal{A},\mathcal{B}) &= \sqcap_{x\in X} \zeta_{\Rightarrow_{TF}}(\mathcal{A}(x),\mathcal{B}(x)) \\ &= \sqcap_{x\in X} (\mathcal{R}_{\mathfrak{TS}}(\mathcal{A}(x),\mathcal{B}(x)) \overset{\propto}{\wedge} \mathcal{R}^{-1}_{\mathfrak{TS}}(\mathcal{A}(x),\mathcal{B}(x))) \end{aligned} \tag{9}$$

Theorem 1. *Let* $\mathcal{A},\mathcal{B}$ *are two* TFS *on* $X = \{x\}$ *and* $\mathcal{R}_{TF}$ *be a residuated* TF *implication with Łukasieuwicz t-norm. Then*

$$d(\mathcal{A},\mathcal{B}) = \bigoplus \{1^\star - \zeta_{\Rightarrow_{TF}}(\mathcal{A}(x),\mathcal{B}(x))\}$$

is Hamming distance between $\mathcal{A}$ *and* $\mathcal{B}$. *Where*

$$\bigoplus\{a_1,a_2,...,a_n\} := a_1 + a_2 + ... + a_n$$

Corollary 1. *Let* $\mathcal{A} = (t_A,\tilde{T}_A,f_A,\tilde{F}_A)$ *and* $\mathcal{B} = (t_B,\tilde{T}_B,f_B,\tilde{F}_B)$ *be a two* TFS *on arbitrary finite set* X *and* $\mathcal{R}_{TF}$ *be a residuated* TF *implication with Łukasieuwicz t-norm. Then*

$$\begin{aligned} d(\mathcal{A},\mathcal{B}) &= \bigoplus \{1^\star - \zeta_{\Rightarrow_{TF}}(\mathcal{A}(x_i),\mathcal{B}(x_i))\} \\ &= \sum_{i=1}^{n} (|t_A(x_i) - t_B(x_i)| + |\tilde{T}_A(x_i) - \tilde{T}_B(x_i)| \\ &\quad + |f_A(x_i) - f_B(x_i)| + |\tilde{F}_A(x_i) - \tilde{F}_B(x_i)|) \end{aligned}$$

is Hamming distance between $\mathcal{A}$ *and* $\mathcal{B}$. *Where*

$$|\tilde{T}_A(x_i) - \tilde{T}_B(x_i)| := |T_A^-(x_i) - T_B^-(x_i)| + |T_A^+(x_i) - T_B^+(x_i)|$$

Theorem 2. *Let* $\mathcal{A},\mathcal{B}$ *are two* TFS *on* $X = \{x\}$ *and* $\mathcal{R}_{TF}$ *be a residuated* TF *implication with Łukasieuwicz t-norm. Then*

$$d(\mathcal{A},\mathcal{B}) = \max\{1^\star - \zeta_{\Rightarrow_{TF}}(\mathcal{A}(x),\mathcal{B}(x))\}$$

is Hausdorff distance between $\mathcal{A}$ *and* $\mathcal{B}$.

Corollary 2. *Let* $\mathcal{A} = (t_A,\tilde{T}_A,f_A,\tilde{F}_A)$ *and* $\mathcal{B} = (t_B,\tilde{T}_B,f_B,\tilde{F}_B)$ *be a two* TFS *on arbitrary finite set* X *and*

$\mathcal{R}_{TF}$ be a residuated TF implication with Łukasieuwicz t-norm. Then

$$d(\mathcal{A},\mathcal{B}) = \frac{1}{n}\sum_{i=1}^{n}\big\{\max\{1^\star - \zeta_{\Rightarrow_{TF}}(\mathcal{A}(x_i),\mathcal{B}(x_i))\}\big\}$$
$$= \frac{1}{n}\sum_{i=1}^{n}\big\{\max\{|t_A(x_i)-t_B(x_i)|, |\tilde{T}_A(x_i) - \tilde{T}_B(x_i)|, |f_A(x_i)-f_B(x_i)|, |\tilde{F}_A(x_i)-\tilde{F}_B(x_i)|\}\big\}$$

is Hausdorff distance between $\mathcal{A}$ and $\mathcal{B}$.

Theorem 3. *Let $\mathcal{A},\mathcal{B}$ are two TFS on finite set X and $\mathcal{R}_{TF}$ be a residuated TF implication with Łukasieuwicz t-norm. Then*

$$d(\mathcal{A},\mathcal{B}) = \bigoplus\big\{1^\star - \sqcap_i\{\zeta_{\Rightarrow_{TF}}(\mathcal{A}(x_i),\mathcal{B}(x_i))\}\big\}$$

is Chebyshev distance between $\mathcal{A}$ and $\mathcal{B}$.

Corollary 3. *Let $\mathcal{A} = (t_A,\tilde{T}_A,f_A,\tilde{F}_A)$ and $\mathcal{B} = (t_B,\tilde{T}_B,f_B,\tilde{F}_B)$ are two TFS on infinite set X and $\mathcal{R}_{TF}$ be a residuated TF implication with Łukasieuwicz t-norm. Then*

$$d(\mathcal{A},\mathcal{B}) = \bigoplus\big\{1^\star - \sqcap_i\{\zeta_{\Rightarrow_{TF}}(\mathcal{A}(x_i),\mathcal{B}(x_i))\}\big\} = \tilde{\bigoplus_i}|\mathcal{A}(x_i)-\mathcal{B}(x_i)|$$

is Chebyshev distance between $\mathcal{A}$ and $\mathcal{B}$. Where,

$$\tilde{\bigoplus_i}|\mathcal{A}(x_i)-\mathcal{B}(x_i)| = (\sup_i|t_A(x_i)-t_B(x_i)| + \sup_i|\tilde{T}_A(x_i)-\tilde{T}_B(x_i)| + \sup_i|f_A(x_i)-f_B(x_i)| + \sup_i|\tilde{F}_A(x_i)-\tilde{F}_B(x_i)|)$$

Proposition 1. *Let $\mathcal{A}$ and $\mathcal{B}$ are two TFS on infinite set X and $\mathcal{R}_{TF}$ be a residuated TF implication with Łukasieuwicz t-norm. Then $d(\mathcal{A},\mathcal{B}) = 1^\star - \mathcal{SI}_{TF}(\mathcal{A},\mathcal{B})$ is Chebyshev distance between $\mathcal{A}$ and $\mathcal{B}$.*

Theorem 4. *Let $\mathcal{A},\mathcal{B}$ are two TFS on set $X=\{x\}$ and $\mathcal{R}_{TF}$ be a residuated TF implication with Łukasieuwicz t-norm. Then*

$$d(\mathcal{A},\mathcal{B}) = \bigoplus\big\{\tilde{\cup}\{1^\star - \zeta_{\Rightarrow_{TF}}(\mathcal{A}(x),\mathcal{B}(x))\}\big\}$$

is Moore distance between $\mathcal{A}$ and $\mathcal{B}$. Where $\tilde{\cup}$ is union operation on interval-valued intuitionistic fuzzy set.

Theorem 5. *Let $\mathcal{A},\mathcal{B}$ are two TFS on finite set X and $\mathcal{R}_{TF}$ be a residuated TF implication with Łukasieuwicz t-norm. Then*

$$d(\mathcal{A},\mathcal{B}) = \tilde{\cup}\{1^\star - \mathcal{SI}_{TF}(\mathcal{A},\mathcal{B})\}$$

is Moore distance between $\mathcal{A}$ and $\mathcal{B}$.

REFERENCES

[1] Atanassov, K.T., Intuitionistic fuzzy implications and modus ponens, Academy of Sciences, 11(1)(2005), 1–5.

[2] Baczy´nski, M. and Jayaram, B., Fuzzy Implications, Studies in Fuzziness and Soft Computing, vol.231, Springer, Berlin, Heidelberg, 2008.

[3] Baczy´nski, M. and Jayaram, B., (S,N)-and R-implications: a state-of-the-art survey, Fuzzy Sets and Systems, 159, (2008) 1836–1859.

[4] Borzooei, R.A., Mohseni Takallo, M., Jun, Y.B., True-False Sets, 2020 8th Iran. Joi. Cong. on Fuzz. and intell. Syst. (CFIS), Publisher: IEEE.

[5] Borzooei, R. A., Fuzzy Sets Theory, (In persian book).

[6] Chen, M.F., Tzeng, G.H., Combining grey relation and TOPSIS concepts for selecting an expatriate host country, Math. Comput. Modell. 40(13) (2004) 1473–1490.

[7] De Baets, B., Mesiar, R., Triangular norms on product lattices, Fuzzy Sets and Systems, 104(1) (1999), 61–75.

[8] Deschrijver, G., The Archimedean property for t-norms in interval-valued fuzzy set theory, Fuzzy Sets Syst. 157 (2006), 2311–2327.

[9] Deschrijver, G., Generalized arithmetic operators and their relationship to t-norms in interval-valued fuzzy set theory, Fuzzy Sets Syst. 160 (2009), 3080–3102.

[10] Deschrijver, G., Cornelis, C., Representability in interval-valued fuzzy set theory, Int. J. Uncertain. Fuzziness Knowl.-Based Syst. 15 (2007) 345–361.

[11] Deschrijver, G., Cornelis, C., Kerre, E.E., On the representation of intuitionistic fuzzy t-norms and t-conorms, IEEE Trans. Fuzzy Syst. 12 (2004) 45–61.

[12] Euler, L., Solutio problematis ad geometriam situs pertinentis, Commentarii academiae scientiarum Petropolitanae, 8, (1741), 128-140.

[13] Jenei, S., A more efficient method for defining fuzzy connectives, Fuzzy Sets Syst. 90(1) (1997) 25–35.

[14] Klement, E.P., Mesiar, R., Pap, E., Triangular Norms, Springer Book, Netherlands, 2000.

[15] Luo, M., Zhao, R., A distance measure between intuitionistic fuzzy sets and its application in medical diagnosis, Artif. Intel. In Medi., 89, (2018), 34–39.

[16] Menger, K., Statistical Metrics, Proc Natl Acad Sci U S A., 28(12) (1942), 535–537. doi: 10.1073/pnas.28.12.535.

[17] Mohseni Takallo, M., Borzooei, R.A., Jun, Y.B., MBJ-neutrosophic structures and its applications in BCK/BCI-algebras, Neut. set and Sys., 23 (2018), 72–84.

[18] Mohseni Takallo, M., Borzooei, R.A., Jun, Y.B., True-False structures on groups and BCK/BCI-algebra, Submitted.

[19] Mohseni Takallo, M., Borzooei, R.A., Jun, Y.B., True-False ideals of BCK/BCI-algebras, Submitted.

[20] Mohseni Takallo, Aaly Kologani M., Borzooei, R.A., Jun, Y.B., Commutative True-False ideals in BCK-algebras and BCI-algebras, Submitted.

[21] Onat, N.C., Gumus, S., Kucukvar, M., Tatari, O., Application of the TOPSIS and intuitionistic fuzzy set approaches for ranking the life cycle sustainability per-

formance of alternative vehicle technologies, Sustainable Prod. Consumption, 6, (2015), 12–25.
[22] Schweizer, B. and Sklar, A., Probabilistic Metric Spaces, North-Holland, NewYork, 1983.
[23] Schweizer, B. and Sklar, A., Statistical metric spaces, Paeifie J. Math., 10, 313-334.
[24] Serstnev, A.N., On a probabilistic generalization of metric spaces, Kazan. Gos. Univ. Ucen . Zap. 124 (1964), 3–11.
[25] Sunil, M., Sunitha, M. S., Types of arcs in a fuzzy graph, Information Sciences, (2009), 1760-1768.
[26] S. Weber, A general concept of fuzzy connectives, negations and implications based on t-norms and t-conorms, Fuz. Sets and Syst., 11 (1983) 115–134.
[27] Zadeh, L. A., Fuzzy Sets, Information and Control, 8(3), (1965), 338–353.
[28] Zadeh, L. A., The concept of a linguistic variable and its application to approximate reasoning I, Inform. Sci., 8, (1975), 199–249.
[29] Zadeh, L. A., The concept of a linguistic variable and its application to approximate reasoning II, Inform. Sci., 8, (1975), 301–357.
[30] Zadeh, L. A., The concept of a linguistic variable and its application to approximate reasoning III, Inform. Sci., 9, (1975), 43–80.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Weak solutions to fuzzy stochastic differential equations under sub-fractional Brownian motion

1st Hossein Jafari, 2nd M. J. Ebadi and 3rdHamed Farahani

1st Department of Mathematics, *Chabahar Maritime University*, Chabahar, Iran, jafari@cmu.ac.ir
2nd Department of Mathematics, *Chabahar Maritime University*, Chabahar, Iran, ebadi@cmu.ac.ir
3rd Department of Mathematics, *Chabahar Maritime University*, Chabahar, Iran, farahani@cmu.ac.ir

***Abstract*— In this paper, the fuzzy stochastic differential equations (FSDEs) driven by sub-fractional Brownian motion (sfBm) are considered which are applied to describe phenomena subjected to randomness and fuzziness simultaneously. The sfBm is known as an extension of the Bm that preserves numerous attributes of fractional Brownian motion (fBm), but not the stationary of the increments. This property makes sfBm a possible candidate to models involving non-stationary increments, self-similarity, and long-range dependence which is suitable for the construction of stochastic models in finance and non-stationary queueing systems. We introduce an approximation approach to the fractional stochastic integrals, and a decomposition of the sfBm to find the existence and uniqueness of the weak solutions.**

***Index Terms*— Fuzzy sets, Sub-fractional Brownian motion (sfBm), Fuzzy stochastic integral(FSI), Stochastic differential equation(SDE), Picard iteration method.**

I. Introduction

Stochastic differential equations (SDEs) under fuzzy stochastic processes (FSPs) can be utilized in systems that are relevant to two kinds of uncertainty, random and fuzzy simultaneously. For instance, in financial markets, fuzzy theory performs fine because financial markets contain uncertainties which differs from stochastic behavior and it is difficult to identify the exact values. We can consider two approaches in this case. On the one hand, several fuzzy parameters are proposed, like volatility and interest rates, to the financial model ([12], [18]). On the other hand, the process of underlying price follows a process of fuzzy stochastic ([17], [19]). The second method avoids the assumption that the parameter is fuzzy in the model and makes it easy enough to obtain α-cuts of the model. There are different approaches on the study of FSDEs. The definition and solutions of the fuzzy Itô integral are given in [5], and the applications of fuzzy stochastic Itô integral are presented in [9], [10]. The fuzzy stochastic Itô integral is considered as an embedding Itô integral into fuzzy space in this method, in which the fuzzy Itô integral is a fuzzy random variable (FRV). The sfBm can be applied as the driving process instead of Brownian motion in the modeling of the systems. The self similarity, long-range dependence, and non stationary increments makes this process suitable for the construction of stochastic models and the analysis of phenomena. Stochastic integral respecting sfBm with parameter $H \in (0, 1/2)$ was defined in [14]. The Wiener integral with respect to an sfBm with index $H \in (0, 1)$ and its domain was characterized in [15], [16].

In the current study, we introduce an FSDE under an sfBm.Such equations can be useful in systems that contain behaviors like stochastic, fuzzy, long-range dependence structure, and non-stationary. We exploit an approximation method to stochastic integral driven by sfBm to find the solutions. Consider the Liouville form of the fBm with parameter $H \in (\frac{1}{2}, 1)$, and the decomposition of the sub-fractional Brownian motion to find the weak solutions.

The paper is organized as follows. The Liouville form and definition of fBm, and sfBm are given in Section 2. Then, several preliminaries on FRVs, FSPs, and fuzzy stochastic integrals (FSIs) are presented. In Section 3, a class of FSDEs driven by sfBm is considered. We use an approximation approach and Picard's iterative method to prove the existence and uniqueness of the solution. The sum up of the paper is given in Section 4.

II. Preliminaries

A. Fractional processes fBm and sfBm

The fBm with Hurst parameter $H \in (0, 1)$ is a centered Gaussian process $B_H = \{B_H(t), t \in [0, T]\}$ and

$$\begin{aligned} R_{B_H}(t, s) &= \mathbb{E}(B_H(t)B_H(s)) \\ &= \frac{1}{2}(s^{2H} + t^{2H} - |t - s|^{2H}). \end{aligned} \tag{1}$$

For $H > 1/2$, this process has self-similarity and long-range dependence properties make it useful in stochastic models.

The classical Itô integral cannot be defined with respect to this process because the fBm is not a semimartingale for $H \neq 1/2$. One of the solutions for this problem is on the basis of the Malliavin calculus method. We have the following representation [11]:

$$B_H(t) = \frac{1}{\Gamma(1+\lambda)} \left(\int_{-\infty}^{0} [(t-s)^{\lambda} - (-s)^{\lambda}] dW(s) + B^H(t) \right),$$

where W is a Wiener process, $\lambda = H - \frac{1}{2}$, and $B^H(t) = \int_0^t (t-s)^{\lambda} dW(s)$. For $H \in (0,1)$, the process $B^H(t)$ is known as the Liouville form of a fBm that has many properties of a fBm except stationary of increments. The process $B^H(t)$ can be approximated by semimartingales by Malliavin calculus method as follows (see [1]):

$$B^{H,\epsilon}(t) = \int_0^t (t-s+\epsilon)^{\lambda} dW(s), \qquad \epsilon > 0, \tag{2}$$

or

$$B^{H,\epsilon}(t) = \lambda \int_0^t \varphi^{\epsilon}(s) ds + \epsilon^{\lambda} W(t), \tag{3}$$

where

$$\varphi^{\epsilon}(t) = \int_0^t (t-s+\epsilon)^{\lambda-1} dW(s). \tag{4}$$

The process $B^{H,\epsilon}(t)$ converges to $B^H(t)$ in $L^2(\Omega)$ when ϵ tends to 0 (see [13]).

The sfBm has attributes similar to those of fBm: Hölder continuity of order γ for any $\gamma < H$, long-range dependence, and self-similarity. Comparing to fBm, the sfBm has non-stationary increments and the increments over non-overlapping intervals are correlated more weakly and its covariance decreases polynomially at higher rate. We can refer to [3] and [16] for the Wiener integral respecting the subfractional Brownian motion.

Consider the subfractional Brownian motion $(S_t^H)_{t\in[0,T]}$, which is continuous Gaussian process with

$$R_{S^H}(t,s) = s^{2H} + t^{2H} - \frac{1}{2}\left[(s+t)^{2H} + |t-s|^{2H}\right], \tag{5}$$

as covariance, in which $H \in (0,1)$. We can show that

$$\begin{aligned} (t-s)^{2H} &\leq \mathbb{E}\left[(S_t^H - S_s^H)^2\right] \\ &\leq (2 - 2^{2H-1})(t-s)^{2H}, H \in (0, \frac{1}{2}), \end{aligned} \tag{6}$$

$$\begin{aligned} (2 - 2^{2H-1})(t-s)^{2H} &\leq \mathbb{E}\left[(S_t^H - S_s^H)^2\right] \\ &\leq (t-s)^{2H}, H \in (\frac{1}{2}, 1). \end{aligned} \tag{7}$$

From the covariance functions, we have

$$\begin{aligned} R_{S^H}(t,s) &> R_{B^H}(t,s) \quad if \;\; H \in (0, \frac{1}{2}) \\ R_{S^H}(t,s) &< R_{B^H}(t,s) \quad if \;\; H \in (\frac{1}{2}, 1). \end{aligned} \tag{8}$$

Consider the following process

$$X_t^H = \int_0^{\infty} (1 - e^{-\theta t}) \theta^{-H-\frac{1}{2}} dW_{\theta}, \tag{9}$$

where W is a standard Bm. The process $X^H = \{X_t^k,\ t \geq 0\}$ is Gaussian with covariance

$$Cov(X_t^H, X_s^H) = \begin{cases} \frac{\Gamma(1-2H)}{2H}\left[t^{2H} + s^{2H} - (t+s)^{2H}\right] & \text{if } H \in (0, \frac{1}{2}) \\ \frac{\Gamma(2-2H)}{2H(2H-1)}\left[(t+s)^{2H} - t^{2H} - s^{2H}\right] & \text{if } H \in (\frac{1}{2}, 1) \end{cases} \tag{10}$$

Moreover, X^H has a version with trajectories which are absolutely continuous on $[0,\ \infty)$ and infinitely differentiable on $(0,\ \infty)$. In [2] a decomposition of the sfBm was obtained in terms of fBm, and an absolutely continuous process as follows

$$S_t^H \overset{d}{=} B_t^H + C_1(H) X_t^H, \tag{11}$$

for $H \in (0, \frac{1}{2})$, in which $C_1(H) = \sqrt{\frac{H}{\Gamma(1-2H)}}$, the Wiener process W and the sfBm S^H are independent. For $H \in (\frac{1}{2}, 1)$, the following decomposition obtained

$$B_t^H \overset{d}{=} S_t^H + C_2(H) X_t^H, \tag{12}$$

where $C_2(H) = \sqrt{\frac{H(2H-1)}{\Gamma(2-2H)}}$, the Wiener process W and the fBm B^H are independent.

B. Fuzzy set theory

Some results and definitions on fuzzy set theory and FSIs are reviewed in this Section. One can refer to [6]–[10] and references therein for more information. Let us denote by $\mathcal{V}(R)$ the family of all non-empty, convex and compact subsets of real numbers R. Denote by d_H, the Hausdorff measure that is given by

$$d_H(A,B) = max\left\{ \sup_{a\in A} \inf_{b\in B} |a-b|, \sup_{b\in B} \inf_{a\in A} |a-b| \right\}.$$

The space $(\mathcal{V}(R), d_H)$ is a separable complete metric space. Let (Ω, Σ, P) be a probability space. The mapping $F : \Omega \to \mathcal{V}(R)$ is $\Sigma-$ measurable if $\{\omega \in \Omega : F(\omega) \cap C \neq \phi\} \in \Sigma$, for any closed set $C \subset R$. Let $\mathcal{M}(\Omega, \Sigma; \mathcal{V}(R))$ be a family of $\Sigma-$ measurable multifunctions with values in $\mathcal{V}(R)$. A multifunction $F \in \mathcal{M}(\Omega, \Sigma; \mathcal{V}(R))$ is L^p- integrably bounded, for $p \geq 1$, if there exists $h \in L^p(\Omega, \Sigma, P; R_+)$ such that $|||F||| \leq h\ p-a.e$, where $R_+ = [0, \infty)$, and

$$|||F||| = d_H(F, \{0\}) = \sup_{f\in F} \|f\|, \qquad for \;\; F \in \mathcal{V}(R),$$

From [4] we know that $F \in \mathcal{M}(\Omega, \Sigma; \mathcal{V}(R))$ is L^p-integrably bounded, for $p \geq 1$, iff $|||F||| \in L^p(\Omega, \Sigma, P; R_+)$. Denote that,

$$\begin{aligned} &\mathcal{L}^p(\Omega, \Sigma, P; \mathcal{V}(R)) \\ &= \{F \in \mathcal{M}(\Omega, \Sigma; \mathcal{V}(R)) : |||F||| \in L^p(\Omega, \Sigma, P; R_+)\}. \end{aligned}$$

For a fuzzy set $Z \in R$, the membership function is defrined by $Z : R \to [0,1]$, where $Z(x)$ is the membership degree of x in Z. Denote by $\mathcal{F}(R)$ the fuzzy sets $Z : R \to [0,1]$ such that $[Z]^{\alpha} \in \mathcal{V}(R)$ for every $\alpha \in [0,1]$, where $[Z]^{\alpha} = \{x \in R : Z(x) \geq \alpha\}$.

We define $d_{\infty} : \mathcal{F}(R) \times \mathcal{F}(R) \to [0, \infty)$ by,

$$d_{\infty}(Z_1, Z_2) = \sup_{\alpha\in[0,1]} d_H([Z_1]^{\alpha}, [Z_2]^{\alpha}).$$

Define $\langle 0\rangle \in \mathcal{F}(R)$ as $\langle 0\rangle := \mathbb{1}_{\{0\}}$, where for $y \in R$, $\mathbb{1}_{\{y\}}(x) = 1$ if $x = y$ and $\mathbb{1}_{\{y\}}(x) = 0$ if $x \neq y$.

Definition 1: An FRV is a function $X : \Omega \to \mathcal{F}(R)$, if the mapping $[X]^\alpha : \Omega \to \mathcal{V}(R)$ is an $\Sigma-$measurable multifunction for all $\alpha \in [0, 1]$.

A mapping $X : \Omega \to \mathcal{F}(R)$ is an FRV iff $X : (\Omega, \Sigma) \to (\mathcal{F}(R), \mathcal{B}_{d_s})$ is $\Sigma|\mathcal{B}_{d_s}-$measurable, with the Skorohod metric d_s in $\mathcal{F}(R)$, and $\mathcal{B}_{d_s}$ is the $\sigma-$algebra generated by the topology induced by d_s. A mapping $X : \Omega \to \mathcal{F}(R)$, is L^p-integrably bounded, $p \geq 1$, if $[X]^0 \in \mathcal{L}^p(\Omega, \Sigma, P; \mathcal{V}(R))$. Let $\mathcal{L}^p(\Omega, \Sigma, P; \mathcal{F}(R))$ be the set of all FRVs that are L^p-integrably bounded.

Let $\widetilde{T} := [0, T]$. Consider a complete, filtered probability space $(\Omega, \Sigma, \{\Sigma_t\}_{t\in\widetilde{T}}, P)$ with a filtration $\{\Sigma_t\}_{t\in\widetilde{T}}$ satisfying an increasing and right continuous family of sub $\sigma-$algebras of Σ, and contains all $P-$null sets.

Definition 2: Consider an FRV $X(t) : \Omega \to \mathcal{F}(R)$ as a mapping for each $t \in \widetilde{T}$, $X : \widetilde{T} \times \Omega \to \mathcal{F}(R)$ is known as an FSP.

Definition 3: X as a FSP is $d_\infty-$continuous, in the case that all mappings $X(\cdot, \omega) : \widetilde{T} \times \Omega \to \mathcal{F}(R)$ as its trajectories are $d_\infty-$continuous functions.

Proposition 1: If X is a FSP and $[X]^\alpha : \widetilde{T} \times \Omega \to \mathcal{V}(R)$ is $\mathcal{B}(T) \otimes \Sigma-$ measurable for all $\alpha \in [0, 1]$, in which $\mathcal{B}(I)$ represents the Borel $\sigma-$algebra of subsets of $\widetilde{T}$, then we say X is measurable.

Proposition 2: X as a fuzzy process is called nonanticipating iff for every $\alpha \in [0, 1]$, $[X]^\alpha$ is measurable with regards to the σ-algebra $\mathcal{G}$, that can be defined as below:

$$\mathcal{G} := \left\{ A \in \mathcal{B}(I) \otimes \mathcal{A} : A^t \in \mathcal{A}_t \text{ for every } t \in \widetilde{T} \right\},$$

where $A^t = \{\omega : (t, \omega) \in A\}$.

Definition 4: Consider X as an FSP X, it is called L^p-integrably bounded ($p \geq 1$), if a stochastic process $h \in L^p(\widetilde{T} \times \Omega, \mathcal{G}; R_+)$ exists such that

$$\left\| |[X(t, \omega)]^0| \right\| \leq h(t, \omega),$$

for almost all $(t, \omega) \in \widetilde{T} \times \Omega$.

The set of L^p-integrably bounded nonanticipating FSPs is denoted by $\mathcal{L}^p(\widetilde{T} \times \Omega, \mathcal{G}; \mathcal{F}(R))$, and let $X \in \mathcal{L}^p(\widetilde{T} \times \Omega, \mathcal{G}; \mathcal{F}(R))$. We define the fuzzy integral as below:

$$\int_0^T X(s, \omega)ds,$$

for $\omega \in \Omega\backslash N_x$, where $N_x \in \Sigma$ and $P(N_x) = 0$. For every $\omega \in \Omega\backslash N_x$ and every $\alpha \in [0, 1]$, the Aumann integral $\int_0^T [X(s, \omega)]^\alpha ds$ belongs to $\mathcal{V}(R)$, so an FRV $\int_0^T X(s, \omega)ds$ belongs to $\mathcal{F}(R)$ for every $\omega \in \Omega\backslash N_x$.

Definition 5: The Lebesgue-Aumann integral Fuzzy stochastic of $X \in \mathcal{L}^1(\widetilde{T} \times \Omega, \mathcal{G}; \mathcal{F}(R))$ is defined by

$$L_x(t, \omega) = \begin{cases} \int_0^T \mathbb{1}_{[0,t]}(s) X(s, \omega)ds & \omega \in \Omega\backslash N_x, \\ \langle 0\rangle & \omega \in N_x. \end{cases} \quad (13)$$

Proposition 3: ([9]) For the FSI L_x, the following properties are given:

1) Let $p \geq 1$. If $X \in \mathcal{L}^p(\widetilde{T} \times \Omega, \mathcal{G}; \mathcal{F}(R))$, then $L_x(\cdot, \cdot) \in \mathcal{L}^p(\widetilde{T} \times \Omega, \mathcal{G}; \mathcal{F}(R))$.
2) For $X \in \mathcal{L}^1(\widetilde{T} \times \Omega, \mathcal{G}; \mathcal{F}(R))$, $\{L_x(t)\}_{t\in\widetilde{T}}$ is $d_\infty-$continuous.
3) For $X, Y \in \mathcal{L}^p(\widetilde{T} \times \Omega, \mathcal{G}; \mathcal{F}(R))$, for $p \geq 1$,

$$\sup_{u\in\widetilde{T}} d^p_\infty(L_x(u, \omega), L_y(u, \omega)) \leq t^{p-1} \int_0^t d^p_\infty(X(s), Y(s))ds.$$

An embedding of R into $\mathcal{F}(R)$ denoted by $\langle\cdot\rangle : R \to \mathcal{F}(R)$ i.e.

$$\langle x\rangle (y) = \begin{cases} 1 & \text{for } y = x \\ 0 & \text{for } y \in R\backslash\{x\} \end{cases}.$$

for $y \in R$.

Definition 6: For a random variable $X : \Omega \to R$ defined on (Ω, Σ, P), the embedding $\langle X\rangle : \Omega \to \mathcal{F}(R)$ is an FRV .

To define a FSI, $\left\langle \int_0^T h(s)dW(s) \right\rangle$ is considered as an FRV in which W is a Wiener process. The below property is useful.

Proposition 4: [9] Let $h \in L^2(\widetilde{T} \times \Omega, \mathcal{G}; R)$, then $\left\{ \left\langle \int_0^t h(s)dW(s) \right\rangle \right\}_{t\in I}$ is a FSP and we have $\left\langle \int_0^t h(s)dW(s) \right\rangle \in \mathcal{L}^2(\widetilde{T} \times \Omega, \mathcal{G}; \mathcal{F}(R))$.

III. FSDE DRIVEN BY SUB-FBM

Consider the following class of FSDEs driven by a sfBm.

$$\begin{aligned} X(t) =& X_0 + \int_0^t f_1(s, X(s))ds + \\ & \left\langle \int_0^t f_2(s, X(s))dS^H(s) \right\rangle, X_0 = X(0), \end{aligned} \quad (14)$$

where $X_0 : \Omega \to \mathcal{F}(R)$ is an FRV , $f_1 : \widetilde{T} \times \Omega \times \mathcal{F}(R) \to \mathcal{F}(R)$, $f_2 : \widetilde{T} \times \Omega \times \mathcal{F}(R) \to R$, and B_H is LfBm with $H \in (\frac{1}{2}, 1)$. The following equation is known as corresponding approximation

$$\begin{aligned} X^\epsilon(t) =& X_0 + \int_0^t f_1(s, X^\epsilon(s))ds + \\ & \left\langle \int_0^t f_2(s, X^\epsilon(s))dS^{H,\epsilon}(s) \right\rangle. \end{aligned} \quad (15)$$

Assumption 1: Consider the following assumptions:

H1) $f_1 : \widetilde{T} \times \Omega \times \mathcal{F}(R) \to \mathcal{F}(R)$ is $\mathcal{G} \otimes \mathcal{B}_{d_s}|\mathcal{B}_{d_s}-$measurable, and $f_2 : \widetilde{T} \times \Omega \times \mathcal{F}(R) \to R$ is $\mathcal{G} \otimes \mathcal{B}_{d_s}|\mathcal{B}(R)-$measurable.

H2) There exists $L > 0$ such that a.e. for each $x, y \in \mathcal{F}(R)$ and every $t \in \widetilde{T}$ it yields

$$\max\{d^2_\infty(f_1(t, \omega, x), f_1(t, \omega, y)), |f_2(t, \omega, x) - f_2(t, \omega, y)|^2\} \leq L^2 d^2_\infty(u, v).$$

H3) There exists $C > 0$ such that for every $t \in \widetilde{T}$, a.e. and for every $x \in \mathcal{F}(R)$ it holds

$$\max\{d^2_\infty(f_1(t, \omega, x), \langle 0\rangle), |f_2(t, \omega, x)|^2\} \leq C^2(1 + d^2_\infty(x, \langle 0\rangle)).$$

Proposition 5: [9] Suppose that $X_1, X_2 \in L^2(\widetilde{T} \times \Omega, \mathcal{G}; R)$, then for each $t \in I$

$$\mathbb{E}\sup_{\nu\in[0,t]} d_\infty^2\left(\left\langle\int_0^\nu X_1(s)dW(s)\right\rangle,\left\langle\int_0^\nu X_2(s)dW(s)\right\rangle\right) \leq 4\mathbb{E}\int_0^t d_\infty^2(\langle X_1(s)\rangle,\langle X_2(s)\rangle)ds. \tag{16}$$

Proposition 6: Suppose that $X_1, X_2 \in L^2(\widetilde{T}\times\Omega,\mathcal{G};R)$, then for each $t\in I$

$$\mathbb{E}\sup_{\nu\in[0,t]} d_\infty^2\left(\left\langle\int_0^\nu X_1(s)dX^H(s)\right\rangle,\left\langle\int_0^\nu X_2(s)dX^H(s)\right\rangle\right) \leq C_{T,H}\mathbb{E}\int_0^t d_\infty^2(\langle X_1(s)\rangle,\langle X_2(s)\rangle)ds, \tag{17}$$

where $C_{T,H}=4\left(2-2^{2H}\right)\Gamma(1-2H)T^{2H-1}$.

Theorem 1: Let $X_0\in\mathcal{L}^2(\Omega,\Sigma_0,P;\mathcal{F}(R))$, and suppose that $f_1:\widetilde{T}\times\Omega\times\mathcal{F}(R)\to\mathcal{F}(R)$, $f_2:\widetilde{T}\times\Omega\times\mathcal{F}(R)\to R$, satisfy assumptions (H1)-(H3). Then, the equation(15) has a unique weak solution.

Proposition 7: The solution $X^\epsilon(t)$ of the equation (15) converges to the solution $X(t)$ of the equation (14) in $\mathcal{L}^2(I\times\Omega)$ as $\epsilon\to 0$ uniformly for $t\in[0,T]$.

IV. Conclusions

In this paper, we focused on a fuzzy stochastic differential equation under the sfBm. We applied an approximation method for stochastic integral under the fBm, and a decomposition of the sfBm into fBm and a Gaussian process with trajectories that are infinitely differentiable. We defined the embedding of stochastic integrals under sfBm into fuzzy space, which is the extension of classical Itô integral under Bm. We defined a fuzzy sub-fractional stochastic differential equation an the existence and uniqueness of the solutions were obtained by using Picard iteration method. Also, uniformly convergence of the approximate solution to the exact solution was proved.

References

[1] E. Alòs, O. Mazet, D. Nualart (2000). *Stochastic calculus with respect to fractional Brownian motion with Hurst parameter lesser than 1/2*. Stoch. Proc. Appl. 86, 121-139.

[2] X. Bardina, D. Bascompte (2009). *A decomposition and weak approximation of the subfractional Brownian motion*. Arxiv, 0905-4360.

[3] T. Bojdecki, L. Gorostiza, A. Talarczyk (2004). *Sub-fractional Brownian motion and its relation to occupation times*. Statistics and Probability Letters, 69, 405-419.

[4] F. Hiai, H. Umegaki (1977). *Integrals, conditional expectation, and martingales of multivalued functions*. J. Multivariate Anal. 7, 149-182.

[5] J. H. Kim (2005). *On fuzzy stochastic differential equations*. J. Korean Math. Soc. 42, 153-169.

[6] V. Lakshmikantham, R. N. Mohapatra (2003). *Theory of Fuzzy Differential Equations and Inclusions*. Taylor and Francis, London.

[7] H. Jafari, M. J. Ebadi (2020).*Malliavin calculus in statistical inference: Cramer-Rao lower bound for fuzzy random variables*. Journal of Decisions and Operations Research, 5(2), 124-132, https://doi.org/10.22105/DMOR.2020.234911.1152.

[8] H. Jafari, M. T. Malinowski, M. J. Ebadi (2021). *Fuzzy stochastic differential equations driven by fractional Brownian motion. Advances in Difference Equations, 16. https://doi.org/10.1186/s13662-020-03181-z*

[9] M. T. Malinowski (2012). Strong solutions to stochastic fuzzy differential equations of Itô type. Math. Comput. Modell. 55, 918-928.

[10] M. T. Malinowski (2013). Some properties of strong solutions to stochastic fuzzy differential equations. Information Sciences 252, 62-80.

[11] B. B. Mandelbrot, J. W. Van Ness (1968). Fractional Brownian motions, fractional noises and applications. SIAM Review 10, 422-437.

[12] P. Nowak, M. Romaniuk (2010). Computing option price for Levy process with fuzzy parameters , European J. Oper. Res., 201(1), 206-210.

[13] T. H. Thao (2006). An approximate approach to fractional analysis for finance. Nonlinear analysis: real world applications, 7(1), 124-132.

[14] G. Shen, C. Chen (2012). Stochastic integration with respect to the sub-fractional Brownian motion with $H\in(0,1/2)$. Stat. Probab. Lett. 82(2), 240-251.

[15] C. Tudor (2008). Inner product spaces of integrands associated to sub-fractional Brownian motion. Stat. Probab. Lett. 78, 2201-2209.

[16] C. Tudor (2009). On the Wiener integral with respect to a sub-fractional Brownian motion on an interval. J. Math. Anal. Appl. 351, 456-468

[17] X. Wang, J. He (2006). A geometric Levy model for n-fold compound option pricing in a fuzzy framework, Journal of Computational and Applied Mathematics, 306, 248-264.

[18] H. C. Wu (2004). Pricing European options based on the fuzzy pattern of Black-Scholes formula, Comput. Oper. Res., 31(7), 1069-1081.

[19] Y. Yoshida (2003). The valuation of European options in uncertain environment, European J. Oper. Res., 145(1) , 221-229.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Zero sets in $MV-$algebras of continuous functions

1st Mahta Bedrood, 2nd Farhad Sajadian and 3rd Arsham Borumand Saeid

1st,3rd Department of Pure Mathematics, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman, Kerman, Iran,
bedrood.m@gmail.com,arsham@uk.ac.ir
2nd Department of Pure Mathematics and Calculations , Faculty of Mathematics, Higher Education Complex of Bam, Kerman, Iran,
fsajadian@bam.ac.ir

***Abstract*— In this article, we study $MV-$algebra of continuous functions $C(X)$ and their zero sets in X. In particular, the properties of zero sets of complemented element in $C(X)$ have been investigated. We study interior of zero sets and examine their relationship with minimal prime ideals of $C(X)$. Also, a topological basis for X is introduced by the interior of zero sets. In addition, we obtain the condition that the union of two zero sets is equal to topological space X.**

***Index Terms*— $MV-$algebra, zero set, zero divisor, complemented element.**

I. Introduction and Preliminaries

$MV-$algebras were introduced by C.C. chang to provide algebraic semantics for Łukasiewicz in finite- valued propositional logic [2]. Chang's completeness theorem states that any MV-algebra equation holding in the standard MV-algebra over the interval [0,1] will hold in every MV-algebra. These algebras relate to the above mentioned system of logic in the same manner as Boolean algebras relate to two classical valued logic. Considering any topological space (shortly in the sequel, space) and [0, 1] endowed with the natural topology, the family C(X) of all [0, 1]-valued continuous functions defined on X has a structure of MV-algebra, induced pointwise by the MV -operations on [0, 1]. The same operations induce on $[0,1]^X$, if X is a nonempty set, the MV-algebra of all the fuzzy sets of X, called usually Bold algebra of fuzzy sets of X [1]. In this paper, $C(X)$ is the $MV-$algebra of all continuous function on completely regular space X to standard $MV-$algebra $([0,1],\oplus,*,0)$. For each $f \in C(X)$, the set $Z(f) = \{x \in X : f(x) = 0\}$, is the zero set of f. For $M \subseteq X$, by $intM$ and $\overline{M}$ we mean the interior and the closure of M. If f is a complemented element of $C(X)$, then $Z(f)$ is an open subset of X and if f is an element of $C(X)$ such that $inf(f(X)) \neq 0$ then $Z(f)$ is an open subset of X. Furthermore, every $f \in C(X)$ is not a zero divisor if and only if $intZ(f) \neq \emptyset$. It is proved that for each complemented element e of $C(X)$, $Ann(e) = (e^*]$ and $Z(e)$ is an open subset of X.
We recollect some definitions and results which will be used in the sequel:

Definition 1. *[2] An $MV-$algebra is a structure (A, $\oplus$, *, 0) where $\oplus$ is a binary operation, *, is a unary operation, and 0 is a constant such that the following axioms are satisfied for any $x, y \in A$:*
(MV1) (A, $\oplus$, 0) is an abelian monoid,
(MV2) $(x^*)^* = x$,
(MV3) $0^* \oplus x = 0^*$,
(MV4) $(x^* \oplus y)^* \oplus y = (y^* \oplus x)^* \oplus x$.

Note that we have $1 = 0^*$ and the auxiliary operation $\odot$ which are as follows:

$$x \odot y = (x^* \oplus y^*)^*.$$

We recall that the natural order determines a bounded distributive lattice structure such that $x \vee y = x \oplus (x^* \odot y) = y \oplus (x \odot y^*)$ and $x \wedge y = x \odot (x^* \oplus y) = y \odot (y^* \oplus x)$. Also for any two elements $x, y \in A$, $x \leq y$ if and only if $x^* \oplus y = 1$ if and only if $x \odot y^* = 0$

Lemma 1. *[3] In each $MV-$algebra A, the following relations hold for all $x, y, z \in A$:*
(1) *If $x \leq y$, then $x \oplus z \leq y \oplus z$ and $x \odot z \leq y \odot z, x \wedge z \leqslant y \wedge z$,*
(2) $x, y \leq x \oplus y$ *and* $x \leq nx = x \oplus x \oplus \cdots \oplus x$,
(3) *If $x \leq y$ and $z \leq t$, then $x \oplus z \leq y \oplus t$,*
(4) $x \wedge (y \oplus z) \leqslant (x \wedge y) \oplus (x \wedge z)$; *in particular* $(mx) \wedge (ny) \leqslant mn(x \wedge y)$, *for every* $m, n \geq 0$.

For any $MV-$algebra A we shall denote by $B(A)$ the set of all complemented elements of $L(A)$ such that $L(A)$ is distributive lattice with 0 and 1.
In the paper A is an $MV-$algebra.

Theorem 1. *[6] For every element e in A, the following conditions are equivalent:*
(1) $e \in B(A)$,
(2) $e \vee e^* = 1$,
(3) $e \wedge e^* = 0$,

(4) $e \oplus e = e$,
(5) $e \odot e = e$.

Definition 2. *[3] An ideal of A is a nonempty subset I of A satisfying the following conditions:*
(I1) If $x \in I,\ y \in A$ and $y \leq x$, then $y \in I$,
(I2) If $x, y \in I$, then $x \oplus y \in I$.
We denote by $Id(A)$ the set of all ideals of A.

Definition 3. *[3] Let I be an ideal of an MV-algebra A. Then I is proper, if $I \neq A$.*
• A proper ideal I of an MV-algebra A is called prime ideal if for all $x, y \in A, x \wedge y \in I$, then $x \in I$ or $y \in I$.
We denote the set of all prime ideals of an MV-algebra A, by $Spec(A)$.
• An ideal I of an MV-algebra A is called a minimal prime ideal of A:
$1) I \in Spec(A)$;
$2)\ If\ there\ exists\ Q \in Spec(A)\ such\ that\ Q \subseteq I,\ then\ I = Q.$
We denote the set of all prime minimal ideals of an MV-algebra A by $Min(A)$.

Note: Let $X \subseteq A$. Define

$$V(X) = \{P \in Min(A) : X \subseteq P\}.$$

Definition 4. *[6] Let X be a nonempty subset of A. Then $Ann(X)$ is the annihilator of X defined by:*

$$Ann(X) = \{a \in A : a \wedge x = 0, \forall x \in X\}.$$

Remark 1. *[6] Let $X \subseteq A$. The ideal of A generated by X will be denoted by $(X]$. We have*
$(1)(X] = \{a \in A \mid a \leqslant x_1 \oplus x_2 \oplus ... \oplus x_n,\ for\ some\ n \in \mathbb{N}\ and\ x_1, ..., x_n \in X\}$. *In particular,* $(a] = \{x \in A \mid x \leqslant na,\ for\ some\ n \in \mathbb{N}\}$.

Definition 5. *[4] Let X be a nonempty subset of A. The set of all zero-divisors of X is denoted by $Z_X(A)$ and is defined as follows:*

$$Z_X(A) = \{a \in A : \exists 0 \neq x \in X\ such\ that\ x \wedge a = 0\}.$$

Zero element of an $MV-$algebra is a zero divisor, which is called trivial zero divisor. We denote by Z_A the set of all zero divisors of A.

On can easily show that $Ann(X) \subseteq Z_X(A)$.

Theorem 2. *[4] Let A be an $MV-$algebra and $P \in Spec(A)$. Then $P \in Min(A)$ if and only if for each $x \in P$, there exists $r \in A - P$ such that $x \wedge r = 0$.*

Theorem 3. *[4] Let A be an $MV-$algebra, $P \in Min(A)$ and I is finitely generated ideal. Then $I \subseteq P$ if and only if $Ann_A(I) \nsubseteq P$.*

Lemma 2. *[4] Let A be an $MV-$algebra. If we have $0 \neq x \in A$, then there exists $P \in Min(A)$ such that $x \notin P$.*

Lemma 3. *[5] Let $X = A \cup B$ such that A and B be closed subsets of X. Also, let $f : A \to Y$ and $g : B \to Y$ be continuous functions. If $f(x) = g(x)$ for all $x \in A \cap B$, then there exists continuous function $h : X \to Y$ such that $h(x) = f(x)$ for all $x \in A$, and $h(x) = g(x)$ for all $x \in B$.*

Theorem 4. *[5] Let X be a topological space. If ζ is a collection of compact subsets of X such that every finite intersection of elements ζ be nonempty, then intersection of all the elements of ζ is nonempty.*

Theorem 5. *[5] Let X be a compact space and $f \in C(X)$. Then there exist $c, d \in X$ such that $f(c) \leq f(x) \leq f(d)$, for all $x \in X$.*

II. Zero sets and their interior in X

Let X be a completely regular space. In this paper, we denote by $C(X)$ the $MV-$algebra of all continuous functions on topological space X to standard $MV-$algebra $([0,1], \oplus, *, 0)$. For every $f, g \in C(X)$ we define $(f \oplus g)(x) = f(x) \oplus g(x), f^*(x) = (f(x))^*$ and $0(x) = 0$, for all $x \in X$. Obviously, $(C(X), \oplus, *, 0)$ is an $MV-$algebra.
Let $f \in C(X)$ and I be an ideal of $C(X)$. Define

$$Z(f) = \{x \in X : f(x) = 0\}\ ,\ Z(X) = \{Z(f) : f \in C(X)\}$$

$$Z(I) = \{Z(f) : \forall f \in I\} \quad , \quad Z^\circ(I) = \{intZ(f) : \forall f \in I\}.$$

$$Z^{-1}(Z(I)) = \{f \in C(X) : Z(f) \in Z(I)\}.$$

$$(Z^\circ)^{-1}(Z^\circ(I)) = \{f \in C(X) : intZ(f) \in Z^\circ(I)\}.$$

Lemma 4. *Let $f_1, f_2 \in C(X)$. Then*
(1) $Z(f_1 \oplus f_2) = Z(f_1) \cap Z(f_2)$,
(2) $intZ(f_1 \oplus f_2) = intZ(f_1) \cap intZ(f_2)$.

Proof: (1) It is clear.
(2) If $x \in intZ(f_1) \cap intZ(f_2)$, then there exist open subsets U_1 and U_2 of X such that $x \in U_1 \subseteq intZ(f_1)$ and $x \in U_2 \subseteq intZ(f_2)$. Put $U = U_1 \cap U_2$. Obviously, $U \subseteq intZ(f_1) \cap intZ(f_2)$. Hence $U \subseteq int(Z(f_1) \cap Z(f_2))$, so $U \subseteq Z(f_1) \cap Z(f_2)$. Thus $x \in Z(f_1) \cap Z(f_2)$ then $f_1(x) = f_2(x) = 0$, hence $(f_1 \oplus f_2)(x) = 0$ so $x \in Z(f_1 \oplus f_2)$. On the other hand U is an open subset of X such that $x \in U$ so $x \in U \subseteq Z(f_1 \oplus f_2)$. Then $x \in intZ(f_1 \oplus f_2)$, implies that $intZ(f_1) \cap intZ(f_2) \subseteq intZ(f_1 \oplus f_2)$. Now, if $y \in intZ(f_1 \oplus f_2)$, then $y \in Z(f_1 \oplus f_2)$. So $(f_1 \oplus f_2)(y) = 0$, thus $f_1(y) = f_2(y) = 0$ which implies that $Z(f_1 \oplus f_2) \subseteq Z(f_1)$ and $Z(f_1 \oplus f_2) \subseteq Z(f_2)$. Hence $intZ(f_1 \oplus f_2) \subseteq intZ(f_1)$ and $intZ(f_1 \oplus f_2) \subseteq intZ(f_2)$. So $intZ(f_1 \oplus f_2) \subseteq intZ(f_1) \cap intZ(f_2)$. Therefore $intZ(f_1 \oplus f_2) = intZ(f_1) \cap intZ(f_2)$. ∎

Theorem 6. *Let $\tau = \{intZ(f) : f \in C(X)\}$. Then τ is a topological basis for X.*

Proof: By Lemma 4(2), it is sufficent to show that for an open set U and $x \in U$, there exists $f \in C(X)$ such that $x \in intZ(f) \subseteq U$. If U is an open subset of X and $x \in U$, then there exists $g \in C(X)$ such that $g(X \setminus U) = \{0\}$ and $g(x) = 1$. Put $f = |(g - (1/4)) \wedge 0|$. Obviously,

$$x \in intZ(f) \subseteq Z(f) = \{x \in X : g(x) \geq 1/4\} = g^{-1}([1/4, 1]) \subseteq U.$$

Therefore τ is a basis for X. ∎

Example 1. *Let $X = \mathbb{R}$ and (a, b) be an open interval in $\mathbb{R}$. Put*

$$f(x) = \begin{cases} 1 & x \in (-\infty, a-1] \\ -x+a & x \in (a-1, a) \\ 0 & x \in [a, b] \\ x-b & x \in (b, b+1) \\ 1 & x \in [b, \infty) \end{cases}$$

Obviously, $(a, b) = intZ(f)$. Then $\tau = \{intZ(f) : f \in C(\mathbb{R})\}$ is a basis for standard topology on $\mathbb{R}$.

Lemma 5. *Let I be an ideal of $C(X)$. Then $Z(I)$ is closed under finite intersections and supersets.*

Proof: Let $Z_1, Z_2 \in Z(I)$. Then there exist $f_1, f_2 \in I$ such that $Z_1 = Z(f_1), Z_2 = Z(f_2)$. Hence $f_1 \oplus f_2 \in I$, so $Z(f_1 \oplus f_2) \in Z(I)$. By Lemma 4(1), $Z(f_1) \cap Z(f_2) \in Z(I)$. Let $Z_1 \in Z(I), Z^{'} \in Z(X)$ and $Z_1 \subseteq Z^{'}$. Then there exist $f_1 \in I$ and $f \in C(X)$ such that $Z_1 = Z(f_1)$ and $Z^{'} = Z(f)$. Hence $f_1 \wedge f \in I$, so $Z(f_1 \wedge f) \in Z(I)$. Obviously, $Z(f) = Z(f_1 \wedge f)$ thus $Z^{'} \in Z(I)$. ∎

Proposition 1. *If I is an ideal of $C(X)$, then $Z^{-1}(Z(I))$ is an ideal of $C(X)$. Also $I \subseteq Z^{-1}(Z(I))$.*

Proof: Obviously, $Z^{-1}(Z(I))$ is a nonempty subset of $C(X)$. Let $f, g \in Z^{-1}(Z(I))$. Then $Z(f), Z(g) \in Z(I)$. By Lemma 5, we get that $Z(f) \cap Z(g) \in Z(I)$ thus $Z(f \oplus g) \in Z(I)$, then $f \oplus g \in Z^{-1}(Z(I))$. Let $f \in Z^{-1}(Z(I)), g \in C(X)$ and $g \leq f$. Then $Z(f) \in Z(I)$ and $Z(f) \subseteq Z(g)$. It follows from Lemma 5, that $Z(g) \in Z(I)$ thus $g \in Z^{-1}(Z(I))$. Obviously, $I \subseteq Z^{-1}(Z(I))$. ∎

Example 2. *Let $X = (0, 1)$ and $I = \{f \in C(X) : \lim_{x \to 0^+} f(x) = 0\}$. Obviously, I is an ideal of $C(X)$. Now, let $g(x) = sin(\frac{\pi x}{2})$, for all $x \in (0, 1)$. Then $Z(g) = \emptyset$, we deduce $\emptyset \in Z(I)$. Also, if $i(x) = 1$, for all $x \in (0, 1)$, then $z(i) = \emptyset$. Hence $i \in Z^{-1}(Z(I))$, so $Z^{-1}(Z(I)) = C(X)$. We obtain $I \subsetneq Z^{-1}(Z(I))$.*

Remark 2. *Let I be an ideal in $C(X)$. It is clear that $Z^{-1}(Z(I)) \subseteq (Z^{\circ})^{-1}(Z^{\circ}(I))$. The reverse inclusion does not hold, in general. For example we consider the ideal $I = \{f \in C(\mathbb{R}) : [0, 1] \cup \{2\} \subseteq Z(f)\}$ in $C(\mathbb{R})$. Assume that $Z(g) = [0, 1]$ and $Z(f) = [0, 1] \cup \{2\}$. Obviously, $intZ(f) = intZ(g)$ and $f \in I$, so $g \in (Z^{\circ})^{-1}(Z^{\circ}(I))$. On the other hand $2 \notin Z(g)$, hence $Z(g) \notin Z(I)$, thus $g \notin Z^{-1}(Z(I))$.*

Proposition 2. *Let $f, g \in C(X)$. Then $intZ(f) = intZ(g)$ if and only if $Ann(f) = Ann(g)$.*

Proof: Let $intZ(f) \subseteq intZ(g)$ and $h \in Ann(f)$. Then $(h \wedge f)(x) = 0$, for all $x \in X$ which implies that $h(x) = 0$ or $f(x) = 0$, so $X \setminus Z(h) \subseteq Z(f)$. Since $Z(h)$ is closed subset of X we get $int(X \setminus Z(h)) = Z(h)$. Hence

$$X \setminus Z(h) \subseteq intZ(f) \subseteq intZ(g) \subseteq Z(g).$$

Then $(g \wedge h)(x) = 0$, for all $x \in X$. Therefore $h \in Ann(g)$. Conversely, let $Ann(f) \subseteq Ann(g)$. To prove that $intZ(f) \subseteq intZ(g)$, it suffices to show that $intZ(f) \subseteq Z(g)$. Suppose $x \in intZ(f)$ and $x \notin Z(g)$. Since $x \notin X \setminus intZ(f)$, then there exists $0 \neq h \in C(X)$ such that $h(X \setminus intZ(f)) = \{0\}$ and $h(x) = 1$. Clearly, $(h \wedge f)(x) = 0$ and $(h \wedge g)(x) \neq 0$, which is impossible. ∎

Theorem 7. *If $f \in C(X)$, then $Ann(f) = \{0\}$ if and only if $intZ(f) = \emptyset$.*

Proof: (1) Let $i \in C(X)$ be such that $i(x) = 1$, for all $x \in X$. Obviously, $Ann(i) = \{0\}, Z(i) = \emptyset$ and $intZ(i) = \emptyset$. Now, if $Ann(f) = \{0\}$, then $Ann(f) = Ann(i)$. It follows from Proposition 1, that $intZ(f) = intZ(i)$. Hence $intZ(f) = \emptyset$.
Conversely, if $intZ(f) = \emptyset$, then $intZ(f) = intZ(i)$. It follows from Proposition 1, that $Ann(f) = Ann(i)$ implies that $Ann(f) = \{0\}$. ∎

Corollary 1. *Let $f \in C(X)$. Then f is not a zero divisor if and only if $intZ(f) = \emptyset$.*

Theorem 8. *Let $f \in A$. Then $f \in B(A)$ if and only if $Ann(f) = (f^*]$.*

Proof: If $f \in B(A)$, then $f \wedge f^* = 0$ so $f^* \in Ann(f)$ hence $(f^*] \subseteq Ann(f)$. Let $x \in Ann(f)$. So $x \wedge f = 0$ thus $x \odot f = 0$, imply that $x \leq f^*$ then $x \in (f^*]$. Hence $Ann(f) \subseteq (f^*]$, therefore $Ann(f) = (f^*]$. Converse is clear. ∎

Theorem 9. *If $e \in B(C(X))$, then $Z(e)$ is an open subset of X.*

Proof: By hypothesis $e \in B(C(X))$, so $(e \oplus e)(x) = e(x)$, for all $x \in X$. We deduce that $e(x) \oplus e(x) = min\{2e(x), 1\} = e(x)$, for all $x \in X$. Hence $e(x) = 0$ or $e(x) = 1$, for all $x \in X$. Put $K = \{x : e(x) = 1\}$. Obviously, $Z(e) \cap K = \emptyset$ and $Z(e) \cup K = X$, then K and $Z(e)$ are clopen subsets of X. Therefore $intZ(e) = Z(e)$. ∎

Theorem 10. *Let $f \in C(X)$ be such that $inf(f(X \setminus Z(f))) \neq 0$ and $Z(f)$ be an open subset of X. Then there exists $e \in B(C(X))$ such that $(e] = (f]$.*

Proof: Define $e : X \to [0, 1]$ by $e(x) = 0$, for all $x \in Z(f)$ and $e(x) = 1$, for all $x \notin Z(f)$. By Lemma 3, we get that $e \in C(X)$. Obviously, $e \in B(C(X))$ and $Z(f) = intZ(f) = Z(e)$. It is claimed that $(f] = (e]$. we consider two cases:
Case 1. if $x \in Z(f)$, then $e(x) = f(x) = 0$. So $e(x) \leq f(x)$.
Case 2. if $x \notin Z(f)$, then $f(x) \neq 0$. By hypothesis $inf f(x) \neq 0$, we imply that there exists $n \in \mathbb{N}$ such that $nf(x) = 1$. So $e(x) \leq nf(x)$.
Hence $e \in (f]$, thus $(e] \subseteq (f]$. Now, it is clear that $(f] \subseteq (e]$. Then $(e] = (f]$. ∎

Proposition 3. *The following statements are equivalent:*
(1) for any zero set Z in X there exists a zero set K in X such that $X = Z \cup K$ and $intZ \cap intK = \emptyset$.
(2) For any $f \in C(X)$ there exists $g \in C(X)$ such that $\overline{(int(Z(f))} = X \setminus Z(g)$.

Proof: $(1 \Rightarrow 2)$ Let $f \in C(X)$ and there exists $g \in C(X)$ such that $X = Z(f) \cup Z(g)$ and $intZ(f) \cap intZ(g) = \emptyset$. Then

$\overline{intZ(f)} \subseteq x \setminus intZ(g)$. On the other hand $\overline{(intZ(g))^c} = \overline{(Z^c(g))}$. Now, it is climed that $\overline{(Z^c(g))} \subseteq \overline{intZ(f)}$. Let $x \in Z^c(g)$. Then there exists a open subset u_x of X such that $x \in u_x$ and $u_x \subseteq Z^c(g)$. Hence $u_x \subseteq Z(f)$, so $x \in \overline{intZ(f)}$. Thus $\overline{(Z^c(g))} \subseteq \overline{intZ(f)}$. so $\overline{intZ(f)} \subseteq Z^c(g) \subseteq \overline{intZ(f)}$. Hence $\overline{intZ(f)} = \overline{Z^c(g)}$.
$(2 \Rightarrow 1)$ Let for any $f \in C(X)$ there exists $g \in C(X)$ such that $\overline{(int(Z(f))} = \overline{X \setminus Z(g)}$. Then $\overline{intZ(f)} = (intZ(g))^c$, so $intZ(f) \cap intZ(g) = \emptyset$. On the other hand $X = Z(f) \cup \overline{(Z(f))^c}$, hence $X \subseteq Z(f) \cup (\overline{intZ(f)})^c$. So $X \subseteq Z(f) \cup (X \setminus \overline{intZ(f)})$, thus $X \subseteq Z(f) \cup intZ(g)$. Hence $X \subseteq Z(f) \cup Z(g)$. Therefore $X = Z(f) \cup Z(g)$. ■

Example 3. *Let $f, g \in C(\mathbb{R})$ such that*

$$f(x) = \begin{cases} 0 & x \in (-\infty, 0) \\ x & x \in [0, 1] \\ 1 & x \in (1, \infty) \end{cases}$$

and

$$g(x) = \begin{cases} 1 & x \in (-\infty, -1) \\ -x & x \in [-1, 0] \\ 0 & x \in (0, \infty) \end{cases}$$

Then $Z(f) = (-\infty, 0]$ and $Z(g) = [0, \infty)$. Obviously, $\mathbb{R} = Z(f) \cup Z(g), intZ(f) \cap intZ(g) = \emptyset$ and $\overline{(int(Z(f))} = (-\infty, 0] = \overline{X \setminus Z(g)}$.

Lemma 6. *Let $f, g \in C(X)$. Then $X = Z(f) \cup Z(g)$ if and only if $g \in Ann(f)$.*

Proof: It is clear. ■

Theorem 11. *Let $f, g \in C(X)$. Then $X = Z(f) \cup Z(g)$ if and only if $V(Ann(f)) \subseteq V(g)$.*

Proof: Let $X = Z(f) \cup Z(g)$ and $P \in V(Ann(f))$. Then $Ann(f) \subseteq P$ so by Lemma 6, $g \in P$. Hence $P \in V(g)$. Conversely, let $V(Ann(f)) \subseteq V(g)$ and $X \neq Z(f) \cup Z(g)$. It follows from Lemma 6, that $g \notin Ann(f)$. Hence there exists $x \in X$ such that $(g \wedge f)(x) \neq 0$. So by Lemma 2, there exists $P \in Min(C(X))$ such that $g \wedge f \notin P$. Thus $f \notin P$ and $g \notin P$. Now, It follows from Theorem 3, that $Ann(f) \subseteq P$, we obtain $P \in V(Ann(f))$. We deduce that $P \in V(g)$, then $g \in P$, which is a contradiction.

■

Lemma 7. *Let $f, g \in C(X)$. Then the following statements are equivalent:*
(1) $V(g) \subseteq V(Ann(f))$;
(2) $V(f) \cap V(g) = \emptyset$;
(3) $V(f \oplus g) = \emptyset$;
(4) $Ann(f \oplus g) = \{0\}$.

Proof: $(1 \Leftrightarrow 2)$ It follows from Theorem 3.
$(2 \Leftrightarrow 3)$ Let $V(f \oplus g) = \emptyset$ and $P \in V(f) \cap V(g)$. Then $f, g \in P$, so $f \oplus g \in P$. Hence $P \in V(f \oplus g)$, which is a contradiction. Conversely, let $V(f) \cap V(g) = \emptyset$ and $P \in V(f \oplus g)$. Then $f \oplus g \in P$, so $f \in P$ and $g \in P$. Thus $P \in V(f) \cap V(g)$, which is a contradiction.
$(3 \Leftrightarrow 4)$ Let $V(f \oplus g) = \emptyset$ and $Ann(f \oplus g) \neq 0$. Then there exists $0 \neq k \in C(X)$ such that $(k \wedge (f \oplus g))(x) = \{0\}$, for all $x \in X$. It followes from Lemma 2, that there exists $P \in Min(C(X))$ such that $k \notin P$. Obviously, $k \wedge (f \oplus g) \in P$, so $(f \oplus g) \in P$. we obtain $V(f \oplus g) \neq \emptyset$, which is a contradiction. Conversely, let $Ann(f \oplus g) = 0$ and $V(f \oplus g) \neq \emptyset$. Then there exists $P \in Min(C(X))$ such that $f \oplus g \in P$. It follows from Theorem 2, that there exists $k \in C(X) \setminus P$ such that $(k \wedge (f \oplus g))(x) = 0$, for all $x \in X$. Obviously, $k \neq 0$ and $k \in Ann(f \oplus g)$, which is a contradiction. ■

Theorem 12. *Let $f, g \in C(X)$. Then $int[Z(f) \cap Z(g)] = \emptyset$ if and only if $V(g) \subseteq V(Ann(f))$.*

Proof: By Lemma 4, Theorem 7 and Lemma7, we have $int[Z(f) \oplus Z(g)] = int[Z(f) \cap Z(g)] = \emptyset$ if and only if $Ann(f \oplus g) = 0$ if and only if $V(g) \subseteq V(Ann(f))$. ■

CONCLUSION

we investigated zero sets of continuous functions $C(X)$ and concluded that this sets and their interior are closed under finite intersections. Also, the set of all interior of zero sets is a topological basis for X. In particular, for every ideal I of $C(X)$ is proved that $Z^{-1}(Z(I))$ is an ideal containing I of $C(X)$. We demonstrated with an example that $Z^{-1}(Z(I))$ is not necessarily equal to I. We proved that zero set of a complemented element is clopen. In addition, for every element f of $C(X)$ is not a zero divisor if and only if $intZ(f) = \emptyset$.

REFERENCES

[1] L. Belluce, " Semisimple algebras of infinite-valued logic and bold fuzzy set theory", Canad. J. Math. 38 (1986), 1356-1379.
[2] C. C. Chang, "Algebraic analysis of many valued logic", Trans. Amer. Math. Soc., **88** (1958), 467-490.
[3] R. Cignoli, I. M. L. D·Ottaviano, D. Mundici, "Algebraic foundations of many- valued reasoning," Kluwer Academic, Dordrecht, (2000).
[4] F. Forouzesh, F. Sajadian, M. Bedrood, "Some results in types of extensions of $MV-$algebras," Publications de l'lnstitut Mathematique, (2019), 161-177.
[5] J. R. Munkres, "Topology", Dorling Kindersely, India (2000).
[6] D. Piciu, "Algebras of fuzzy logic", Ed. Universitaria Craiova (2007).

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Asymmetric distributions based on the t-copula

1st Fereshteh Arad, 2nd Ayyub Sheikhi

1st and 2nd Department of Statistics, *Faculty of Mathematics and Computer*, Shahid Bahonar University of Kerman, Kerman, Iran,
fereshteh_arad@yahoo.com, sheikhy.a@uk.ac.ir.

***Abstract*— Assuming that $C_{X,Y}$ is the copula function of X and Y with marginal distribution functions $F_X(x)$ and $F_Y(y)$, in this work we study the selection distribution $Z \overset{\mathrm{d}}{=} (X|Y > \mu_y)$. We present some special cases of our proposed distribution, among them, skew-t distribution. The simulation results showed that the distribution presented in this paper works better according to the measurement criteria.**

***Index Terms*— copula, selection distribution, skew-t, t-copula.**

I. Introduction

A general univariate form of asymmetric distribution, which is called skew-symmetric family, can be represented as a function of a cumulative distribution function and a density distribution function [8], a random variable Z, whose density function is as follows, is called skew-symmetric.

$$f(z) = 2f_0(z)G(w(z)), \qquad z \in \mathbb{R}, \tag{1}$$

where $f_0(.)$ is a probability function centrally symmetric about 0, $w(.) : \mathbb{R} \to \mathbb{R}$ is an odd real-valued function such that $w(-x) = -w(x)$ and $G(w(.))$ is a cumulative distribution function on $\mathbb{R}$ such that $g = G'$ is an even density function, providing G is differentiable. It is said that the univariate random variable Z_{st} has a skew-t distribution, if its density can be written as

$$f_Z(z) = 2t_d(z;\nu)T_1(\alpha^T w^{-1}(z-\xi)\big(\frac{\nu+d}{Q_z+\nu}\big)^{1/2};\nu+d), \tag{2}$$

where $Q_z = (z-\xi)^T\Omega^{-1}(z-\xi)$ and t_d is the probability density function of a d-dimensional t distribution with ν degrees of freedom, and accordingly $T_1(x;\nu+d)$ denotes the t distribution function with $\nu+d$ degrees of freedom [1]. The aim of this work is the study of conditional distribution of two copula related random variables X and Y. Based on the Sklar theorem [5] for any random vector (X,Y), there exists a grounded, uniformly marginal and $2-$increasing bivariate copula function $C : [0,1]^2 \to [0,1]$ such that

$$F_{X,Y}(x,y) = C(F_X(x), F_Y(y)), \tag{3}$$

where $F_{X,Y} : \mathbb{R}^2 \to [0,1]$ is the joint cumulative distribution function of X, Y and $F_X,\ F_Y : \mathbb{R} \to [0,1]$ are respectively cumulative distribution functions of random variables X and Y. We assume that both X, Y are continuous random variables which guaranties that copula C is unique and hence its density, exists and is just function $c(.,.)$ such that

$$f_{X,Y}(x,y) = f_X(x)f_Y(y)c(F_X(x), F_Y(y)), \tag{4}$$

The relation is easily obtained by deriving the relation (3). We refer to [7, 10] for more information about copulas and association measures. As Resconi and Boris Kovalerchuk 2017 have pointed out, copula is known a bridge between probability theory and fuzzy logic. So, there are many work in the literature related to the connection of copulas and fuzzy logic as well as expert systems, among others, we refer to [6, 9, 11].

This study seeks to calculate $Z \overset{\mathrm{d}}{=} (X|Y \in C)$ when the variables X and Y are related through the copula function. A special case of the copula function is the t-copula. Denoting T the cumulative t-distribution with ν degree freedom and $\boldsymbol{T}(.,\boldsymbol{R})$ the bivariate t-distribution function with correlation matrix $\boldsymbol{R}$, it is said that random variables X and Y are associated to a t-copula with correlation matrix $\boldsymbol{R}$ if their coupling function is

$$C^t(u,v) = \boldsymbol{T}_{\boldsymbol{R},\nu}(T_\nu^{-1}(u), T_\nu^{-1}(v),\ \boldsymbol{R}) \tag{5}$$

where $T_\nu^{-1}(.)$ is the inverse function of the t-distribution function with ν degree freedom and $\boldsymbol{R} \in [-1,1]^{2\times 2}$ is a correlation matrix matrix with off diagonals ρ [6]. We consider this copula as a connection function of our variables in this work.

II. Main results

It is well known that conditioning a variable in a subset of another variable cause of a skewness parameter. Assuming that two random variables X and Y are associated with the

copula function $C_{X,Y}$, then from the classical probability, we have the distribution of $Z \stackrel{\mathrm{d}}{=} (X|Y \in C)$ as

$$\begin{aligned} F_Z(z) &= \frac{\int_{-\infty}^{z}\int_C f_{X,Y}(x,y)dydx}{\int_C f_Y(y)dy} \\ &= \frac{\int_{-\infty}^{x}\int_C f_X(x)f_Y(y)c(F_X(x),F_Y(y))dydx}{\int_C f_Y(y)dy} \\ &= m_C\int_{-\infty}^{x}\int_C f_X(x)f_Y(y)c(F_X(x),F_Y(y))dydx, \end{aligned}$$

where $m_C = (\int_C f_Y(y)dy)^{-1}$ and differentiating with respect to z we readily obtain the density of z as

$$f_Z(z) = m_C\int_C f_X(z)f_Y(y)c(F_X(z),F_Y(y))dy, \qquad z \in \mathbb{R}, \tag{6}$$

where $c(.,.)$ is the density copula.

The following theorem states a more specific version of (6) in which the selection is made on the positive values of Y which are greater than its mean.

Theorem 1: Assume that two continuous random variables X and Y are connected via the copula function $C_{X,Y}$, then the distribution of $Z \stackrel{\mathrm{d}}{=} (X|Y > \mu_y)$ is given by

$$f_Z(z) = m_C f_X(z)(1 - D_1C_{X,Y}), \qquad z \in \mathbb{R}, \tag{7}$$

where $D_1C_{X,Y} = \frac{\partial C_{X,Y}}{\partial F_X(z)}$ if exist otherwise 0.

We have

$$\begin{aligned} F_Z(z) &= m_C P(X \le z, y > \mu_y) \\ &= m_C[P(X \le z) - P(X \le z, y \le \mu_y)] \\ &= 2[F_X(z) - C_{X,Y}(F_X(z), F_Y(\mu_y))] \end{aligned}$$

and using some chain rules we readily obtain

$$\begin{aligned} f_Z(z) &= 2\big[f_X(z) - \frac{\partial C_{X,Y}}{\partial F_X(z)}\frac{\partial F_X(z)}{\partial z}\big] \\ &= 2\big[f_X(z) - f_X(z)D_1C_{X,Y}\big] \\ &= m_C f_X(z)(1 - D_1C_{X,Y}), \end{aligned}$$

which proves the assertion [3]. Recently, we sought to find in the univariate and multivariate mode that the variables are related to each other through the truncated gaussian copula and compared the obtained distributions with the skew normal distribution [2, 3]. In this study we examine the other states that the variables are related through t-copula. The following corollary consider the t-copula in Theorem 1.

Corollary 1: Let X and Y follow t-distribution with ν degrees of freedom and are related by the t-copula with correlation ρ (5), then the distribution of $Z \stackrel{\mathrm{d}}{=} (X|Y > \mu_y)$ is

$$f_Z(z) = 2t_1(z;\nu)T_1\Big(\frac{\rho z}{\sqrt{1-\rho^2}}\big(\frac{\nu+1}{\nu+z^2}\big)^{\frac{1}{2}};\nu+1\Big) \tag{8}$$

Since $m_{\mu_y} = 2$ and $\mu_y = 0$, similar to the proof of Theorem 1 we have

$$F_Z(z) = 2[T_1(z;\nu) - C_{X,Y}(T(z), T(\mu_y))].$$

Regarding to differentiating with respect to z, we first note that

$$\frac{\partial C_{X,Y}(u,v,\rho)}{\partial u} = T_1\Big(\frac{\rho z}{\sqrt{1-\rho^2}}\big(\frac{\nu+1}{\nu+z^2}\big)^{\frac{1}{2}};\nu+1\Big)$$

and again using some chain rules we have

$$\begin{aligned} f_Z(z) &= 2\big[t_1(z;\nu) - \frac{\partial C_{X,Y}}{\partial T(z)}\frac{\partial T(z)}{\partial z}\big] \\ &= 2t_1(z;\nu)T_1\Big(\frac{\rho z}{\sqrt{1-\rho^2}}\big(\frac{\nu+1}{\nu+z^2}\big)^{\frac{1}{2}};\nu+1\Big) \end{aligned}$$

which is 8.

We many consider the marginal distributions do not follow the t-distribution the following corollary provide such situation.

Corollary 2: Let X be a t random variables and Y has a distribution function $F_Y(y)$ with mean μ_y and they are associated with the t-copula with correlation ρ then the distribution of $Z \stackrel{\mathrm{d}}{=} (X|Y > \mu_y)$ is

$$f_Z(z) = m_{\mu_y}t_1(z;\nu)T_1\Big(\frac{\rho z - T^{-1}(F_Y(\mu_y))}{\sqrt{1-\rho^2}}\big(\frac{\nu+1}{\nu+z^2}\big)^{\frac{1}{2}};\nu+1\Big) \tag{9}$$

where $m_{\mu_y} = (\int_{\mu_y} f_Y(y)dy)^{-1}$, The proof is straightforward similar to the proof of Corollary 1 and is omitted.

A. Numerical Results

1) Simulation Study: In order to assess and visualize our proposed distribution, in the section we used a Monte Carlo simulation study. We generated 1000 random pairs $(X_i, Y_i), i = 1, 2, ..., 1000$, in such a way that X followed a t distribution with 10 degree of freedom and Y came from an exponential distribution with $\lambda = 10$ and they are connected via a t-copula with correlation $\rho_{XY} = 0.6$ and $\rho_{XY} = 0.9$ [1]. We then repeated this procedure 5000 times. Tables II summarized the AIC and BIC of our proposed skew-t-copula distribution against the skew-t ones. As seen from this table, skew-t-copula outperforms skew-t, specially for large value of ρ.

TABLE I
AIC AND BIC OF SKEW-T-COPULA AND SKEW-T.

Estimation	skew-t-copula	skew-t
$AIC_{\rho=0.6}$	1153	1154
$BIC_{\rho=0.6}$	1164	1165
$AIC_{\rho=0.9}$	960	971
$BIC_{\rho=0.9}$	962	972

2) Real Data: In this research, (AIS) data has been used in the GLMsData package of R software. In this dataset, variables Ht and WBC follow the t distribution with 2 degree of freedom and exponential with mean 78, respectively. These variables are also associated with t-copula with $\rho = 0.25$. In the following table, the two distributions of skew-t and skew-t-copula are compared. As this table reveals, implementing the associated copula of these two variable yields improvements of AIC and BIC .

[1]We use the R software and all codes are available upon request.

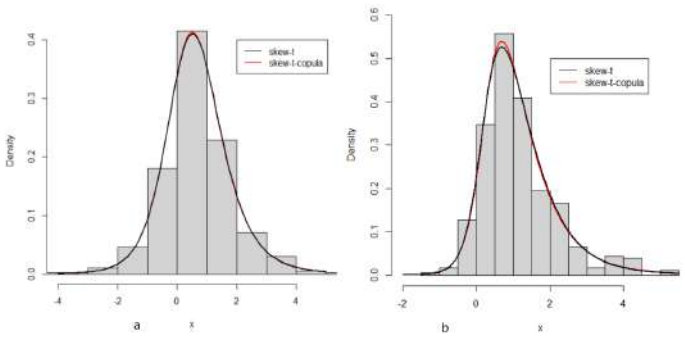


Fig. 1. Performance of skew-t and skew-t-copula, a) $\rho = 0.6$, b) $\rho = 0.9$.

TABLE II
AIC AND BIC OF SKEW-T-COPULA AND SKEW-T.

Estimation	skew-t-copula	skew-t
$AIC_{\rho=0.25}$	43	44
$BIC_{\rho=0.25}$	55	55

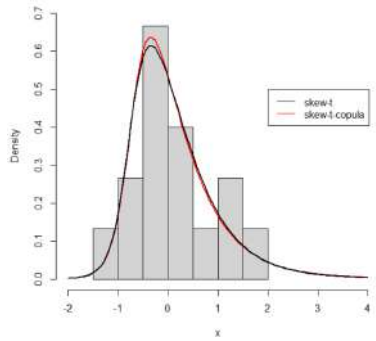


Fig. 2. Performance of skew-t and skew-t-copula, $\rho = 0.25$.

B. Conclusion

The purpose of this paper was to obtain $Z \stackrel{\mathrm{d}}{=} (X|Y > \mu_y)$ such that the variables are related through the t-copula. And the skew-t-copula formula was introduced, and according to the simulated and real data results, according to these hypotheses, if we use skew-t-copula, compared to skew-t, more accurate results will be brought. It can also be explored for future research with the actual data of the given formula.

III. REFERENCES

REFERENCES

[1] A. Azzalini, A. Capitanio,"Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution,"Statistical Methodology, J. Royal Statistical Society, series B, 65, 2003, pp. 367-389.

[2] A. Sheikhi, F. Arad and R. Mesiar,"Multivariate asymmetric distributions based on copula related random variables," Stat, Papers, To appear.

[3] A. Sheikhi, F. Arad and R. Mesiar ,"On asymmetric distribution of copula related random variables which includes the normal ones," Kybernetika, To appear.

[4] A. Sheikhi, V. Amirzade and R. Mesiar,"A comprehensive family of copulas to model bivariate random noise and perturbation," Fuzzy Sets and Systems, 415, 2020, pp. 27-36.

[5] A. Sklar,"Fonctions de reṕartition a n dimensions et leurs marges," Publications de l′Institut Statistique de l′Université de Paris, 8, 1959, pp. 229–231.

[6] B. Kovalerchuk,"Relationships between probability and possibility theories," In Uncertainty Modeling, Springer, Cham, 2017, pp. 97-122.

[7] F. Durante, C. Sempi, Principles of copula theory, CRC press, 2015.

[8] J. Wang, J. Boyer and M.G, Genton,"A skew-symmetric representation of multivariate distributions,"Statistica Sinica, JSTOR, 2004, pp.1259–1270.

[9] P. Hajek, R. Mesiar,"On copulas, quasicopulas and fuzzy logic,"Soft Computing, 2008, 12(2), pp.1239–1243.

[10] R.B. Nelsen, An introduction to copulas, Springer Series in Statistics, Second Edition, Springer-Verlag, New York, 2006.

[11] S. Giakoumakis, B. Papadopoulos,"An algorithm for fuzzy negations based-intuitionistics fuzzy copula aggregation operators in multiple attribute decision making,"Algorithms, 2020, 13(6).

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Improving the fit for diagonal copula based on Kendall's tau and tail dependence preserving transformation

Selim Orhun Susam

Department of Econometrics, *Munzur University*, Tunceli, Turkey, orhunsusam@munzur.edu.tr

***Abstract*—This study proposed the method for improving the goodness of fit for diagonal copula utilizing Kendall?s tau and tail dependence preserving transformation implemented on the diagonal section of the diagonal copula. We investigate the performance of our methodology on real data examples. The findings indicate that the model fitting is improved by using τ and tail dependence preserving transformation.**

***Index Terms*—Diagonal copula, tail dependence, Kendall's tau**

I. Introduction

A copula is a multivariate distribution function that allows us to model dependencies among random variables. The diagonal copula is the most preferred copula constructed from the diagonal section. Since different diagonals yield different diagonal copulas, the collection of diagonal copulas is very large (Nelsen and Fredricks [4]). But, most of the diagonal sections are usually one-parametric models. However, researchers want to model dependencies through diagonal copulas, it would be more rational to work with multi-parametric diagonal sections. Main purpose of this paper is to develop the multi-parameter diagonal section which preserves τ and tail dependence of diagonal copulas utilizing polynomial transformation applied to the existing diagonal section. In addition, we study on the improvement of the fit of diagonal copulas by using these transforms.

There are some contributions in improving the goodness of fit for a copula model. Michiels and De Schepper [3] introduced concordance invariant transforms and studied the improvement the fit of Archimedean copulas by using these transforms. Also, Susam and Ucer [5] studied the τ-preserving transforms to obtain the best fitting multi-parameter Archimedean copula through the goodness-of-fit testing procedure for energy data.

The paper is organized as follows: In Section 2, we discuss some basic properties of copulas with a focus on the diagonal copula class. In section 3, we define τ and tail preserving transformation for diagonal copula under some conditions and give some properties of this transformation based on its diagonal section. In Section 4, on a real data example, we give the goodness- of fit improvement procedure using τ and tail preserved transformed diagonal copula and compare the results numerically and visually and the last section is devoted to the conclusion.

II. Basic concepts

Let X and Y be the random variable having the joint distribution function H and the marginals F and G, respectively. The copula C is the function that links the multivariate joint distribution function to its marginal distributions due to the following relationship proposed by Sklar [6]:

$$H(x, y) = C(F(x), G(y)).$$

Copula C is unique if and only if marginals F and G are continuous. Also, it satisfies the following properties

1) $C(0, u) = C(u, 0) = 0$ for all $u \in [0, 1]$
2) $C(1, u) = C(u, 1) = u$ for all $u \in [0, 1]$
3) for all $u,\ u',\ v,\ v' \in [0, 1]$ with $u < u'$ and $v < v'$

$$V_C([u, u'] \times [v, v']) = C(u, u') - C(u, v') - C(u', v) + C(u, v) \geq 0$$

where $V_C([u, u'] \times [v, v'])$ is the $C - volume$ of the rectangle $[u, u'] \times [v, v']$.

In this paper we will mainly focus on a diagonal copula families. This class is mainly characterized by a diagonal section. $\delta_C : [0, 1] \to [0, 1]$, called diagonal section of copula, is the function defined by $\delta(t) = C(t, t)$. Let us consider that X and Y are uniformly distributed on the unit interval. Moreover, suppose that $W = \max(X, Y)$ is distributed according to the cumulative distribution function (cdf) H. The behaviour of

the random variable W is determined by the diagonal section of the copula $C_{X,Y}$, such that $\delta_C(t) = H_W(t)$ (Durante et al. [1]). Diagonal section of the copula has the following properties:

(D1) $\delta_C(0) = 0$ and $\delta_C(1) = 1$;
(D2) $\delta_C(t) \leq t$ for all $t \in [0,1]$;
(D3) $\delta_C(t)$ is non-decrasing;
(D4) δ_C is $2-Lipschitz$, such that $\big|\delta_C(t_2) - \delta_C(t_1)\big| \leq 2\big|t_2 - t_1\big|$ for all $t_2, t_1 \in [0,1]$.

For the diagonal section, there are some copulas with diagonal which coincide with δ. Let us suppose that δ is a diagonal function satisfied D11-D14. Nelen and Fredrick [4] introduced diagonal copula $K_\delta(y,z)$ with diagonal section δ. They defined $K_\delta(u,v)$ given by

$$K_\delta(u,v) = \min\big(u, v, \frac{\delta(u)}{2} + \frac{\delta(v)}{2}\big),$$

for all $(u,v) \in [0,1]^2$. Kendall's tau for diagonal copula K_δ is given by

$$\tau_K = 4\int_0^1 \delta(t)dt - 1.$$

Also, upper tail dependence (λ_U) and lower tail dependence (λ_L) of diagonal copula can be defined as following:

$$\lambda_U = \lim_{t\to 1^-} \frac{1 - 2t + \delta(t)}{1-t} = 2 - \delta'(1^-).$$

$$\lambda_L = \lim_{t\to 0^+} \frac{\delta(t)}{t} = \delta'(0^+).$$

It is obvious that, there are clear links beetwen dependence coefficents $(\tau,\ \lambda_U,\ \lambda_L)$ and diagonal sections of K_δ. Hence, we can create a new diagonal copula which preserves the dependence coefficients, utilizing polynomial transfromation applied to existing diagonal section.

Now, we present to procedure of estimation of diagonal section. Let $\{(X_1,Y_1),\ldots,(X_n,Y_n)\}$ be a random sample of (X,Y) from cdf $H(x,y)$. The inference is then based on the pseudo-samples defined as

$$U_i = \frac{R(X_i)}{n},\ V_i = \frac{R(Y_i)}{n},\ i = 1,\ldots,n;$$

where $R(.)$ is the rank of random variable. Hence, the pair of random variables (U,V) yield an approximate sample from the copula $C(u,v)$. The non-parametric estimation of diagonal section relies on the pseudo-observations

$$w_i = \max\big(u_i, v_i\big),\ i = 1,\ldots,n \tag{1}$$

which have the distribution function $C(w,w)$. It is natural to non-parametric estimate the diagonal section given by

$$\delta_n(t) = \frac{1}{n}\sum_{i=0}^{n} \mathbf{I}\big(w_i \leq t\big),\ t \in [0,1]; \tag{2}$$

which by the Glivenko-Cantelli lemma converges to the true cdf. Erderly [2] investigated properties of empirical diagonal section δ_n.

III. Kendall's tau and tail dependence preserving transformation

In this section, we investigate a transform that preserves the both value of the Kendall's tau and tail dependence of the the diagonal with diagonal section δ.

Lemma 1. *Let f_α be a continuous increasing function with $f'(0) = 1$ and it has a bounded first and second derivatives on the unit interval. Then $\delta_T(t) = f'_\alpha(t)\delta\big(f_\alpha(t)\big)$ is the Kendall's tau preserving transformation.*

Proof. Let $\tau = 4\int_0^1 \delta_T dt - 1 = 4\int_0^1 f'_\alpha(t)\delta\big(f_\alpha(t)\big)dt - 1$. A substitution $u = f_\alpha(t)$ leads to the desired result. $\delta_T(t)$ satisfies the conditions D1-D4. It is obvious that $\delta_T(0) = 0$ and $\delta_T(1) = 1$ because f_α is increasing and $f_\alpha(0) = 0$, $f_\alpha(1) = 1$, $f'(0) = 1$. We note that, if f_α is a Lipschitz function then Lipschitz constant can be defined as $L_{f_\alpha} = \sup_{t\in[0,1]} \big|f'_\alpha(t)\big|$. In this manner, it has a bounded first derivative. On the other hand, if the followig inequality is satisfied then $\delta_T(t)$ is 2-Lipschitz:

$$\sup_{t\in[0,1]} \big|\delta'_T(t)\big| = \sup_{t\in[0,1]} \big|f''_\alpha(t)\delta\big(f_\alpha(t)\big) + \big(f'_\alpha(t)\big)^2\delta'\big(f_\alpha(t)\big)\big| \leq 2.$$

Clearly, $\delta'\big(f_\alpha(t)\big)$ is bounded because $\delta(t)$ is Lipschitz and $f_\alpha(t)$ is bounded. In addition, $\sup_{t\in[0,1]} \big|\delta'_T(t)\big|$ is bounded if f_α has a bounded first and second derivatives. □

Next lemma gives the tail dependence coefficients of transformed diagonal copula.

Lemma 2. *Kendall's tau preserving transformation given by $\delta_T(t) = f'_\alpha(t)\delta\big(f_\alpha(t)\big)$ has the following tail dependence properties:*

$$\lambda_U^T = \lambda_U - f''(1);$$

$$\lambda_L^T = \lambda_L f'_\alpha(0)^2.$$

If we apply extra conditions on the Kendall's tau preserving transformation $\delta_T(t) = f'_\alpha(t)\delta\big(f_\alpha(t)\big)$ defined in Lemma 1, it enables us to construct transforms which not only preserve the Kendall's tau, but also the tail dependence coefficients.

Lemma 3. *Let f_α be a continuous increasing function with $f''(1) = 0$, $f'_\alpha(0) = 1$ and $f''(0)$ bounded. The transform given by $\delta_T(t) = f'_\alpha(t)\delta\big(f_\alpha(t)\big)$ also preserves the tail dependence.*

IV. Real data application

In this section, we illustirate the use of Kendall's tau and the tail dependence preserving transformation for diagonal copulas in order to increase its goodness-of-fit. For this purpose, we use data set which includes population density in people per square mile and Fuel consumption in 10.000.000 US gallons per year in 26 cities of U.S. The data set which is called as "road" can be found in the R package "MASS".

In order to assess the goodness-of fit result for diagonal copula, we use Cramér-Von Mises statistic which measures the distance between the empirical diagonal section and the null hypothesis diagonal section given by

$$CvM = \int_0^1 n\Big(\delta_n(t) - \delta_\theta(t)\Big)^2 d\delta_n(t). \tag{3}$$

Thus, the CvM statistic defined in Equation (4.1) allow us to compare the distances among copulas. The p-values of the test statistics which are obtained by running 10.000 Monte Carlo samples by the bootstrap method. All goodness-of-fit results can be found in Table 1.

Let $\delta_{\alpha,\beta}(t) = t^2 + \alpha t(1-t)^2 + \beta t^2(1-t)$, where $\alpha,\ \beta \in [0,1]$; and $K_{\delta_{\alpha,\beta}}(u,v)$ denote the diagonal copula constructed from $\delta_{\alpha,\beta}$. It has the Kendall's tau, lower tail dependence and upper tail dependence coefficients given by respectively

$$\tau = \frac{1}{3}(1+\alpha+\beta),\ \lambda_U = \beta,\ \lambda_L = \alpha,$$

for more details see Nelsen and Fredrick [4]. We start by estimating the Kendall's tau from the data. It is found to be 0.3821 and by inverted Kendall's tau method α and β can be calculated as, respectively, 0.78 and 0.28. In order to asses to fit for $K_{\delta_{\alpha,\beta}}$, CvM test statistic defined in equation (4.1) is found as 0.021 with P-value 0.000. As a result, the copula $K_{\delta_{\alpha,\beta}}$ does not provide an acceptabe fit. The graph of the $\delta_{\alpha,\beta}$ is visulized in Figure 1.

To increase the goodness of fit result, we define a polynomial transformation with order seven as following:

$$f_{T,7}(t) = \sum_{k=1}^{7} \alpha_k t^k.$$

To ensure the properties of diagonal section defined by D1-D4 and conditions for a tau and tail preserving transform, let $w_i,\ i = 1,\dots,26$ be the psuedo observations defined in equation (2.1). Recalling Lemma 1 and 3, the necessary conditions for a tau and tail preserving transform are f_α is increasing, $f_\alpha(0) = 0$, $f_\alpha(1) = 1$, $f'(0) = 1$, $f''(1) = 0$. This conditions implies that $\sum_{k=1}^{7} \alpha_k = 1$, $f'_{T,7}(w_i) \geq 0$, $\alpha_1 = 1$ and $\sum_{k=0}^{7} \alpha_k k(k-1) = 0$. The equality $f_\alpha(0) = 0, f_\alpha(1) = 1$ is neccesseray condition to satisfy property D1. To statisfy the condition D2, we impose that $\delta_{T,7}(w_i) \leq w_i,\ i = 1,\dots,26$. The non-decreasing conditions defined by D3 impose that $\delta'_{T,7}(w_i) \geq 0$ and the Lipschitz conditions D4 are interpreted as $\left|\delta'_{T,7}(w_i)\right| \leq 2,\ i = 1,\dots,26$. The parameters $\alpha_k,\ k = 1,\dots,7$ can be estimated by minimizing the Cvm test statistic defined in equation (4.1) considering the following constraints:

1) $\alpha_1 = 1$,
2) $\sum_{k=1}^{7} \alpha_k = 1$,
3) $f'_{T,7}(w_i) \geq 0$
4) $\sum_{k=0}^{7} \alpha_k k(k-1) = 0$
5) $\delta_{T,7}(w_i) \leq w_i,\ i = 1,\dots,26$,
6) $\delta'_{T,7}(w_i) \geq 0,\ i = 1,\dots,26$,
7) $\left|\delta'_{T,7}(w_i)\right| \leq 2,\ i = 1,\dots,26$.

After applying the constraints above in minimizing problem, the parameters α_k can be considered as a shape parameters. Estimated parameters α_k are represented in Table 1. The general dependence Kendall's tau and dependence in the tails does not change under this transform. Transformed diagoal section $\delta_{T,7}$ is visulized in Figure 1.

Similarly, we can create Kendall's tau and tail dependence preserving transform with nine parameters as following:

$$f_{T,9}(t) = \sum_{k=1}^{9} \alpha_k t^k.$$

We proceed in the same way as with the seven-parametric transform, and estimated parameters are given in Table 1. As a result of this transform CvM test statistic lowers to 0.010 with p-value 0.008. Again, an accaptable fit is not achived. To achieve an acceptable fit, more parameters needs to be added. We use new transform with eleven parameters proceeding in the same way with $f_{T,7}(t)$ and $f_{T,9}(t)$. Now, we have more shape parameters than other transforms. As a result of this transform, CvM test statistic lowers to 0.001 with p-value 0.051. The best fit is achieved for the copula $K_{\delta_{T,11}}$ with the transformed diagonal section $\delta_{T,11}$ which is shown in Figure 1. In Figure 2 the data set is visualized, on the left the observed psuedo-sample, in the center using the $K_{\delta_{\alpha,\beta}}$, and on the right we have the transformed copula with eleven parameters.

TABLE I
GOODNESS-OF-FIT RESULTS FOR K-SC

Copula	$\hat{\theta}$	$\hat{\beta}$	$\hat{\alpha_1}$	$\hat{\alpha_1}$	$\hat{\alpha_3}$	$\hat{\alpha_4}$	$\hat{\alpha_5}$	$\hat{\alpha_6}$	$\hat{\alpha_7}$	$\hat{\alpha_8}$	$\hat{\alpha_9}$	$\hat{\alpha_{10}}$	$\hat{\alpha_{11}}$	CvM	P-Val
$K_{\delta_{\alpha,\beta}}$	0.78	0.28	—	—	—	—	—	—	—	—	—	—	—	0.021	0.000
$K_{\delta_{T,7}}$	0.78	0.28	0	5.14	-7.12	-4.55	20.01	-17.56	5.08	—	—	—	—	0.015	0.002
$K_{\delta_{T,9}}$	0.78	0.28	0	11.69	-70.83	251.71	-532.58	673.54	-497.30	196.90	-32.13	—	—	0.010	0.008
$K_{\delta_{T,11}}$	0.78	0.28	0	7.45	-36.44	161.40	-487.67	810.44	-512.86	-425.62	972.40	-638.01	149.91	0.001	0.051

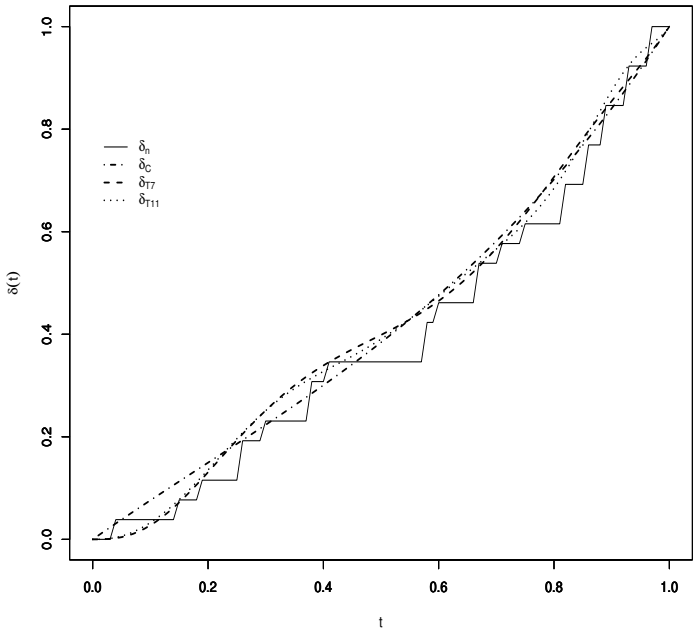
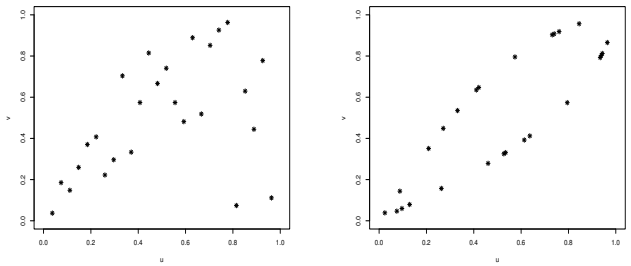
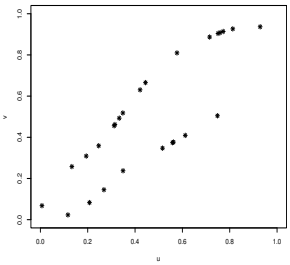


Fig. 1. Graph of the diagonal sections for real data set



(a) Real data set

(b) Diagonal copula with δ_C



(c) Transformed diagonal copula with δ_C for $m = 11$

Fig. 2. Resampling the data set

V. Conclusion

In this study, we investigate the improving fit for diagonal copula with Kendall?s tau and tail dependence preserving transform which maintains membership in the family while increasing the number of shape parameters. The methodology was illustrated by improving the goodness of fit on the data set. Findings indicate that Kendall?s tau and tail dependence preserved transformed copula as better fit to the data.

References

[1] Durante, F., Kolesárová, A., Mesiar, R., and Sempi, C. (2007). Copulas with Given Diagonal Sections: Novel Constructions and Applications. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, **15**(4), 397-410.

[2] Erdely, A. (2008) Exact distribution under independence of the diagonal section of the empirical copula, *Kybernetika*, 44(6), 826?845.

[3] Michiels, F., Koch, I., and De Schepper, A. (2012). How to improve the fit of Archimedean copulas by means of transforms. *Statistical Papers*, **53**, 345-355.

[4] Nelsen, R.B., and Fredricks, G.A. (1997). "Diagonal copulas" in Distributions with given marginals and moment problems (Prague, 1996), *V. Benes and J. Stepan, Eds. Dordrecht: Kluwer Acad. Publ.* 121?128.

[5] Susam, S.O., and Hudaverdi, B.U. (2019). Modeling the dependence structure of CO_2 emissions and energy consumption based on the Archimedean copula approach: The case of the United States. *Energy Sources, Part B: Economics, Planning, and Policy*, **14** (6): 274?89.

[6] Sklar, A. (1959), Fonctions de repartition a n dimensions et leurs marges. *Publications de lInstitut de Statistique de lUniversite de Paris*, 8, 229-231.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Multivariate ageing intensity

1st Hossein-Ali Mohtashami-Borzadaran, 2nd Ali Dolati, 3rd Mohammad Amini and 4th Magdalena Szymkowiak

1st,2nd & 3rd Department of Statistics, Ferdowsi University of Mashhad, Mashhad, Iran.
4th Institute of *Automatic Control and Robotics, Poznan University of Technology,Poznan, Poland.* h.mohtashami@mail.um.ac.ir.

***Abstract*— In reliability, the ageing intensity function (AIF) characterize the intensity of ageing for lifetime variable. Here, he bivariate and multivariate version of this measure is presented via the copula function. Since, the copula function extracts the raw dependence between random variables, the bivariate and multivariate AIF measures are estimated via copula function.**

***Index Terms*—Reliability, Ageing intensity, Lifetime distribution, Failure rate**

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Testing positive quadrant dependence with discrete copulas

Elisa Perrone

Eindhoven University of Technology, Department of Mathematics and Computer, e.perrone@tue.nl

***Abstract*—In this talk, we propose a new test for positive quadrant dependence based on discrete copulas. To define the test statistic, we exploit the geometric representation of discrete copulas as a polytope. We compare the proposed method with existing tests in various simulated scenarios. We conclude by discussing the advantages and limitations of the proposed method. The talk is based on joint work with Fabrizio Durante and Giuseppe.**

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

On quantifying and estimating directed dependence

Wolfgang Trutschnig

University of Salzburg, wolfgang@trutschnig.net

***Abstract*—Considering that a (random) variable X may provide more information about a (random) variable Y than vice versa it is natural that dependence measures, i.e., notions quantifying the extend of dependence, are not necessarily symmetric. Working with Markov kernels (regular conditional distributions) allows to construct the measure ζ_1 which assigns every copula a value in $[0,1]$, which is 0 exactly in the case of independence, and 1 exclusively for Y being a function of X (see [1]), and which can consistently be estimated in full generality (see [2]). Extending ζ_1 to the general multivariate setting of quantifying the extend of dependence of a random variable Y on an ensemble of random variables $X_1,\dots X_d$ is a nontrivial endeavor since different marginal distributions have to be taken into account - working with so-called linkages, however, allows to overcome this problem. The talk will first provide background on ζ_1 and on its checkerboard estimator and then discuss the multivariate approach.**

REFERENCES

[1] W. Trutschnig: On a strong metric on the space of copulas and its induced dependence measure, Journal of Mathematical Analysis and Applications 384, 690-705 (2011), doi:10.1016/j.jmaa.2011.06.013

[2] R.R. Junker, F. Griessenberger, W. Trutschnig: Estimating scale-invariant directed dependence of bivariate distributions, Computational Statistics and Data Analysis 153, 107058 (2021), doi:10.1016/j.csda.2020.107058

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Invariant copulas under univariate truncation

Fabrizio Durante

Dipartimento di Scienze dell'Economia, Università del Salento, Lecce, Italy, fabrizio.durante@unisalento.it

***Abstract*—We consider the class of copulas that are invariant under truncation with respect to one variable. We provide their characterization and present their main properties. Moreover, various results about bounds for this class under special constraints are hence considered, with particular emphasis on the case when either the associated Spearman's rho or Kendall's tau is known.**

Poster Papers

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

A Nitrate Enzymatic Biosensor based on Optimized Machine Learning Techniques

Keyvan Asefpour Vakilian

Department of Biosystems Engineering, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran
keyvan.asefpour@gau.ac.ir, ORCID: 0000-0001-5035-7727

***Abstract*—Many enzymatic biosensing devices have been developed for specific and selective nitrate detection over the last two decades. They generally use the nitrate-nitrite redox reaction to measure the nitrate. Since the activity of the enzyme used in the structure of the biosensor reduces over time, the operator should replace the enzyme immobilized on the working electrode frequently, which increases the detection costs and limits their commercialization capabilities. In this study, artificial neural networks (ANNs) have been used for nitrate concentration prediction in the samples considering both electrochemical data and enzyme activity decrement over time. Genetic algorithm (GA) and particle swarm optimization (PSO) were used to optimize the architecture of the ANNs used for the decision-making of the biosensor. This study showed that both GA-ANN and PSO-ANN learning algorithms resulted in promising nitrate prediction performance with coefficients of determination (R^2) higher than 0.93 and 0.94, respectively. Moreover, the biosensor could be used up to two weeks after the enzyme immobilization for nitrate determination in more than 500 samples after each time of electrode preparation. Finally, a comparison between the findings of this study and former studies that used support vector machines and fuzzy inference systems revealed that ANNs optimized with evolutionary and swarm intelligence techniques can provide more reliable prediction results.**

***Index Terms* —artificial neural networks; genetic algorithm; nitrate detection; particle swarm optimization**

I. INTRODUCTION

The control, monitoring, and removal of environmental pollutants such as nitrate in aquatic ecosystems and plant tissues is essential to reduce its unfavorable effects on the environment and human health [1]. The effective control and removal of nitrate require the reliable determination of its concentration in natural waters, soils, and plants. Particularly, high concentrations of nitrate in drinking water have become a significant concern for humans [2]. Moreover, determining and controlling the nitrate concentration in groundwater and surface water resources is a high priority for maintaining good water quality. International organizations, such as WHO, have determined the concentration of 800 μM as a limit for nitrate in drinking water [3]. Various analytical methods, such as spectrophotometric [4], chromatographic [5], electrochemical [6], electrophoretic [7], flow injection [8], and fluorescence spectroscopy [9], have been utilized for the determination of nitrate. Some limiting requirements, such as special sample pre-treatments, auxiliary reactions for the production of a colored product, and mandatory separation steps in chromatography, reduce the point-of-care and instant detection capabilities in most of these methods [1].

The more reliable techniques for nitrate determination use a biocatalyst for the electrochemical reaction of nitrate in liquid samples [3]. The biocatalyst, which can be *Nitrate reductase* (NaR) or a living organism containing NaR, is used to catalytic reduce nitrate to nitrite. This enzymatic biocatalyst can remarkably improve both electrode sensitivity and selectivity toward nitrate reduction [3]. Various sensors that include selective electrodes for nitrate measurement have been successfully introduced to determine nitrate in a wide range of samples. These sensors that include a bioreceptor for detection purposes are called biosensors [10]. Today, using biosensors for efficient and reliable nitrate concentration determination has gathered remarkable interest as an alternative method for conventional detection methods due to their robustness, simplicity, selectivity, low detection limit, and sensitivity [11].

During the nitrate-nitrite biocatalytic redox reaction performed in a three-electrode system, the biosensor voltammetric response is measured and analyzed by a signal processing unit for nitrate determination [12]. Due to the fact that the activity of the enzyme as bio-macromolecule decreases over time after its immobilization on the working electrode of

the biosensor [13], it is mandatory to periodically remove and replace the enzyme on the electrode to obtain reliable measurement results. For example, the electrochemical signal of a biosensor developed for nitrate detection is reduced by more than 10% even two days after the electrode preparation because of the natural NaR degradation process [14]. The enzyme immobilization and electrode preparation are usually time-consuming and increase the cost of measurements. The requirement for frequent enzyme immobilization undesirably affects the commercialization of biosensors developed for environmental monitoring [13].

To untie this knot, the lifespan of the prepared working electrode along with the electrochemical data of the biosensor can be analyzed to obtain a reliable prediction. In this situation, a smart biosensing device can be developed using artificial intelligence as the decision-making unit [15]. The input variables of the system are the electrochemical data recorded by the biosensor as well as the operating properties of the electrode, while the output variable is nitrate concentration. So far, several intelligent biosensors have been developed for nitrate determination using support vector machines (SVMs) and fuzzy inference systems (FISs) [16-18]. It seems that by developing artificial neural networks (ANNs) with optimized architecture using sophisticated metaheuristic algorithms, more reliable prediction results can be achieved for intelligent biosensors. Therefore, in this study, a supervised machine learning method with characteristics optimized by efficient optimization methods was introduced to consider the effects of decrement of the enzyme activity after the enzyme immobilization on the biosensor electrochemical results. To do this, a three-electrode electrochemical biosensor is developed to record the cyclic voltammetry of the samples in liquid samples, while a decision-making unit is used for nitrate determination based on the electrochemical data.

II. Material and Methods

Figure 1. depicts the schematics of the proposed system. It consisted of two main parts, one for extracting the electrochemical response of the samples under analysis and another for the intelligent decision-making task.

A. Electrochemical Nitrate Biosensing Unit

As depicted in Figure 1, a three-electrode unstirred electrochemical unit containing a glassy carbon (GC) working electrode, a platinum auxiliary electrode, and an Ag/AgCl reference electrode was considered in this study. The detail of the implementation of the biosensor is brought in our previous study [18]. Briefly, 30 mM NaR in 50 mM assay buffer consisting of Bis-Tris-acetate, potassium chloride, magnesium acetate, and calcium chloride was pipetted onto the GC working electrode to prepare the GC/NaR working electrode. The purification method of *Arabidopsis thaliana* NaR, which was immobilized on the working electrode from a heterogeneous expression system in *E. coli*, is described in the literature [19,20]. Anthraquinone-2-sulfonate (AQ) was used during experiments as the mediator compound for electron transfer between the enzyme and the working electrode.

Cyclic voltammetry was carried out utilizing a commercial potentiostat (PalmSens, The Netherlands) to collect the features (i.e., the electrochemical response) of the samples. During the measurements, the potential was applied to the GC working electrode and Ag/AgCl reference electrode, while the electrical current was measured and recorded between the working and platinum electrodes. Voltammetric sweep range was [0 – -400] mV vs. NHE with a sweeping rate of 5 mV/s.

B. Machine Learning Unit

A program written in the MATLAB programming environment was used for machine learning based on the electrochemical response of the samples. The program included the ANN regression model for nitrate concentration prediction in the samples having two sets of input features.

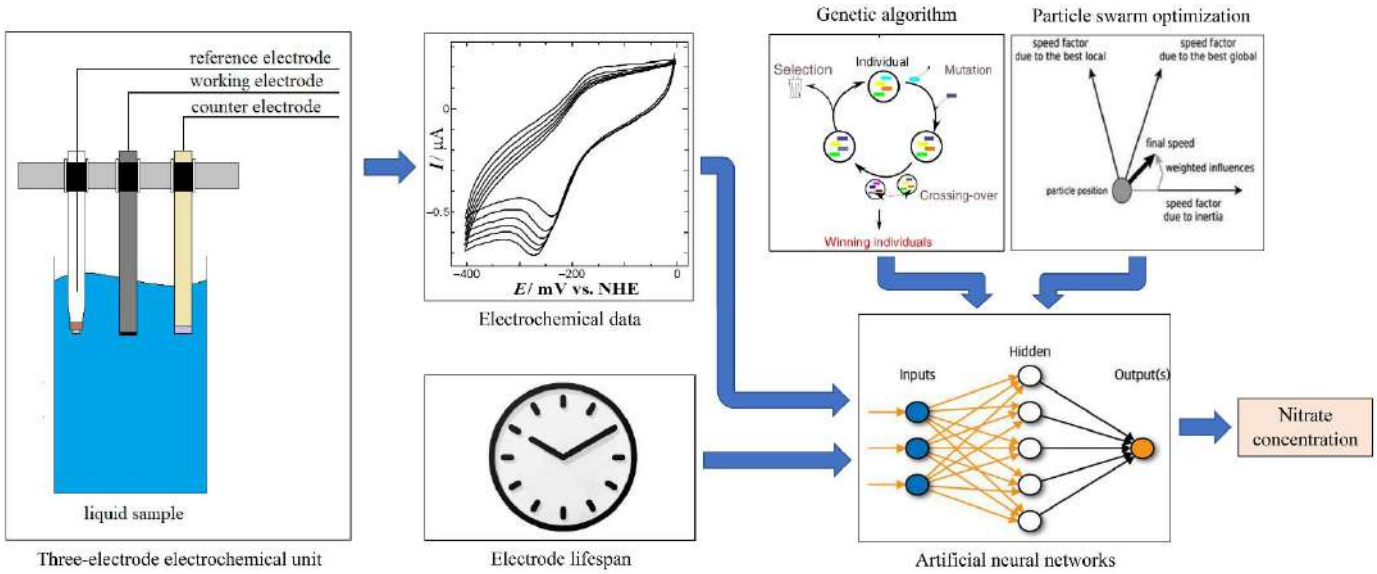


Figure 1. Schematic diagram of the intelligent nitrate biosensor

These input features included the electrochemical data of the sample and the lifespan of the prepared GC/NaR electrode. The number of hidden layers (between 1 and 3) and the number of neurons in each layer (between 5 and 20), and their weight and bias parameters were optimized using metaheuristic optimization algorithms to find the most reliable ANN architecture in predicting nitrate concentration in the samples under analysis (Figure 1). In this study, genetic algorithm (GA) as an evolutionary method and particle swarm optimization (PSO) as an optimization method based on swarm intelligence were used to optimize the architecture of the ANNs.

GA works based on creating an initial set of random populations as possible solutions. Each individual is a chromosome, describing a solution for the problem. GA selects individuals with higher eligibility, being more likely to survive and crossover [21]. After an iterative procedure called generations, the parent chromosomes produce better offspring by removing weak solutions based on the objective function. Like GA, firstly, PSO considers a random population, each component of which is a different set of decision variables whose optimal values are obtained based on an iterative algorithm. A velocity vector and a position vector force the population to change their positions in the search space toward the optimal solutions. The velocity consists of two components: the best position that a particle has ever reached and the best position that other particles in its neighborhood have ever reached. Each particle provides a solution in each iteration. The best result in the population during the iterations is considered as the optimized value [22]. Table 1 shows the parameters of the optimization methods used in this study. The objective function of these methods was to minimize the mean squared error of the prediction using the ANNs. The stop condition during the optimization was to reach the maximum number of iterations.

TABLE I. PARAMETERS OF THE OPTIMIZATION ALGORITHMS

Optimization method	Parameter	Value
GA	Population size	100
	Maximum number of generations	500
	Mutation rate	0.1
	Crossover percentage	0.5
PSO	Initial inertia weight	1
	Inertia weight damping ratio	0.99
	Cognitive acceleration (C_1)	1
	Social acceleration (C_2)	2
	Population size	200
	Maximum number of iterations	100

A dataset was created for training ANN structures to provide a reliable model for recognizing the sample of unknown. To prepare the dataset, aliquots of $NaNO_3$ were added to the inert electrolyte solution to achieve nitrate standards with nitrate contents varying from 0 to 2000 mM. The prepared standards were analyzed by the developed biosensor from immediately after the enzyme immobilization to 20 days after the immobilization.

In this study, 5-fold cross-validation was used to evaluate the performance of the biosensor. Nitrate concentration of the samples as their targets were normalized using Eq. (1)

$$x_n = \frac{x_0 - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

where x_n is the value after normalization, x_0 is the raw value, and $x_{\min}$ and $x_{\max}$ are the lowest and highest nitrate concentration values in the dataset, respectively. Mean squared error (MSE) (Eq. 2) and coefficient of determination (R^2) (Eq. 3) were considered as error evaluation criteria

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(x_\mathrm{p} - x_\mathrm{o})^2 \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(x_\mathrm{p} - x_\mathrm{o})^2}{\sum_{i=1}^{n}(x_\mathrm{o} - \bar{x}_\mathrm{o})^2} \quad (3)$$

where x_o is the concentration of nitrate standards, x_p is the predicted value using ANNs, and n is the number of samples. The lower MSE and higher R^2, the better performance of the machine learning model is. Using these criteria obtained for normalized data, it was possible to compare the performance of the machine learning techniques in the proposed smart biosensing device.

III. RESULTS AND DISCUSSION

The prediction performance of the ANNs optimized by GA and PSO, which are called GA-ANN and PSO-ANN, as well as an ANN with a single learner layer with 10 neurons, is shown in Figure 2 over time. Since the R^2 values are calculated considering the normalized data, the results of the models are comparable. Figure 2 indicates that, in general, the performance of PSO-ANN was better than that of the GA-ANN during the experiments. Using the GA and PSO optimization models, ANN was capable of predicting nitrate content in the samples with R^2 values higher than 0.93 and 0.94 even two weeks after the electrode preparation. However, a non-optimal ANN resulted in R^2 values lower than 0.9 six days after the electrode preparation. Similar to the results of this study, many studies have shown the capability of metaheuristic optimization algorithms in improving the performance of ANNs in regression problems [23-25].

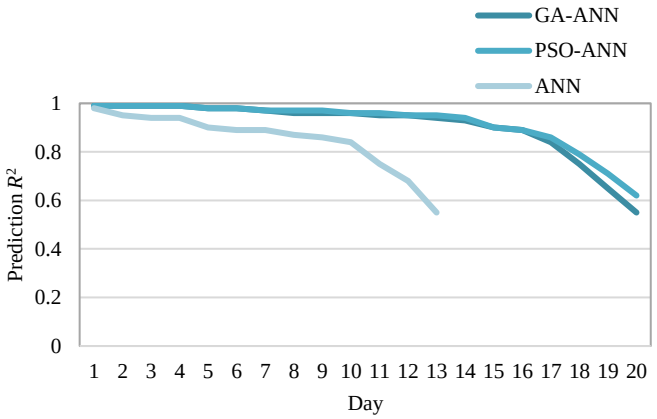


Figure 2. R^2 of the models in predicting the nitrate concentration of samples analyzed during 20 days after enzyme immobilization

According to Figure 2, in the initial days of the experiment, although both optimal models had acceptable performance in the nitrate concentration prediction, ANN with a non-optimal architecture resulted in poor performance. The reason was the complex relationships between the inputs and the output.

Because the introduced biosensor can provide promising results during the 14 days after the electrode preparation based on an $R^2 > 0.9$ threshold, the average MSE and R^2 values during this period were calculated. The obtained values for MSE and R^2 were approximately 0.0004 and 0.97 for both optimized models, respectively.

Similar to other sensors, performance characteristics of the biosensor, i.e., linear range and detection limit, should be reported to compare the performance of the biosensors. In this situation, the biosensor is usually used without delay after the electrode preparation (i.e., enzyme immobilization on the surface of the GC working electrode). The electrochemical data of the samples recorded by the voltammetric unit of the biosensor might be nonlinear during the experiments when we talk about intelligent biosensors. However, the decision-making unit is able to predict the nitrate content with acceptable efficiency. Therefore, similar to other intelligent systems, the performance of the introduced biosensor is evaluated using 5-fold cross-validation.

The main finding of this study relies on providing a machine learning technique to increase the durability of the prepared working electrode of the biosensor. This is essential in environmental monitoring, especially when it is impossible to transfer the liquid samples to a laboratory for nitrate determination. This happens when it is decided to measure the nitrate concentration of a lake or a river. Moreover, the device introduced in this study is cost-effective because it does not require frequent electrode preparation.

During the initial hours after enzyme immobilization on the working electrode, the electrochemical characteristics of the samples were similar to the findings reported in the literature [14]. The electrochemical behavior of the biosensing unit in this short period describes the catalytic reduction reaction of nitrate to nitrite in the presence of a mediator. However, after several days, the electrochemical response of the biosensor was changed entirely because of the remarkable reduction of enzyme activity. This emphasizes the role of machine learning models in improving the performance of an analytical method.

The catalytic activity of most enzymes extracted from living organisms decreases significantly over time [26], and the voltammetric results change remarkably, and thus, they cannot be used for analytical purposes. The reproducibility of a sensor is another performance characteristic showing the performance of the sensor in exerting similar results for liquid samples with similar analyte concentrations. Its quantitative value for the intelligent biosensor can be obtained based on the performance of the decision-making unit, which was acceptable in this study ($R^2 > 0.90$) even two weeks after the immobilization of NaR on the GC working electrode.

The repeatability of the biosensor was evaluated by point-of-care determination of nitrate in arbitrary river water, tap water, and mineral water samples with five replications. Table 2 reveals the mean and standard deviation of the measurements using the introduced intelligent biosensor, which used PSO-ANN to improve its architecture. Table 2 also indicates the nitrate content measured by a conventional spectrophotometric technique [27]. In this laboratory technique, acid-catalyzed nitration of salicylic acid yields a colored product which is further analyzed with UV-Vis spectroscopy. As can be seen in Table 2, the nitrate contents obtained by the proposed device are similar to those obtained by conventional laboratory techniques, which are time-consuming and require expensive instruments. According to the results, the relative standard deviation of the results of the proposed biosensing device has experienced an increase due to the increase in GC/NaR electrode lifespan. However, the standard deviations of the measurements by the proposed smart biosensing device were still lower than those obtained from the conventional laboratory method, indicating the acceptable efficiency of the proposed intelligent biosensor.

TABLE II. A COMPARISON OF THE PROPOSED SYSTEM AND A CONVENTIONAL SPECTROSCOPIC METHOD IN NITRATE DETERMINATION

Sample	**Electrode lifespan (days)**	**Nitrate concentration (μM)**			
		Proposed biosensor based on PSO-ANN		**Spectroscopic method**	
		Mean	**S.D.***	**Mean**	**S.D.***
River water	2	238.0	3.3	240.9	10.3
River water	14	242.7	7.9	240.9	10.3
Tap water	2	37.5	0.7	37.1	2.9
Tap water	14	36.8	1.5	37.1	2.9
Mineral water	2	89.1	2.0	88.5	3.8
Mineral water	14	89.5	2.9	88.5	3.8

* Standard deviation

Table 2 reveals that both biosensing and conventional spectroscopic methods showed higher nitrate concentrations in the river water samples than tap and mineral waters. It was predictable since the nitrogen-based fertilizers can leach from agricultural fields to water resources, causing undesirable consequences for the health of humans and other living organisms. Using the introduced biosensor will be easier to provide contamination maps than before since it will be unnecessary to transfer the unknown liquid samples from the sampling location to the laboratory for analysis. In this situation, all the data can be gathered on-site, which is suitable for point-of-care measurements. Today, intelligent biosensors have also been used to analyze nutrient deficiency and other stresses in plants [28].

Table 3 shows a comparison of intelligent biosensors in predicting nitrate concentration. According to the table, the electrode's durability prepared in the introduced intelligent biosensor is higher than those reported in the literature. This helps analysts to use this device for point-of-care detection of nitrate contamination in the environment. Moreover, although ANN has been used in these devices before [16], the biosensor was equipped with a portable spectroscopic unit along with the electrochemical three-electrode unit, which increased the cost of the device remarkably. FIS decision-making unit is also reported for such sensors [17].

However, they require an expert to define fuzzy rules and appropriate determination of fuzzy membership functions for the model inputs and output. The intelligent biosensors can be helpful to emerge the next generation of portable environmental biosensing devices that can exhibit a selective and specific response to the analyte under study. It is suggested to use nanoparticles, such as gold and silver nanoparticles and nanostars, to improve the biocatalytic activity of the enzyme [29-31] or develop more reliable machine learning methods to improve the analytical and prediction performance of these biosensors.

TABLE III. COMPARISON OF INTELLIGENT BIOSENSORS IN PREDICTING NITRATE CONCENTRATION

Biosensing data	**Machine learning method**	**Durability of the prepared electrode (days)**	R^2	**Reference**
Electrochemical and spectroscopic data	NB	10	0.90	[16]
	DT	10	0.90	[16]
	RF	10	0.84	[16]
	ANN	10	0.89	[16]
	SVM	10	0.96	[16]
	LS-SVM	10	0.92	[16]
Electrochemical data	FIS	-	0.95	[17]
	SVM	10	0.93	[18]
	GA-ANN	14	0.97	This study
	PSO-ANN	14	0.97	This study

NB: Naïve Bayes, DT: decision tree, RF: random forest, LS-SVM: least squares-SVM

IV. CONCLUSION

In this study, the ANN regression model optimized with metaheuristic methods was used for nitrate concentration prediction in a three-electrode electrochemical biosensor. Results indicated that the trained machine could predict analyte content in the samples with promising performance. PSO-ANN resulted in reliable nitrate prediction accuracy during two weeks after the preparation of the GC/NaR electrode with MSE and R^2 of 0.97 and 0.0004, respectively. The findings of this study can be used to develop an emerging generation of point-of-care and cost-effective biosensors for efficient analytical tasks.

REFERENCES

[1] M. E. E. Alahi and S. C. Mukhopadhyay, "Detection methods of nitrate in water: A review," sensor. Actuat. A-Phys., vol. 280, pp. 210–221, 2018.

[2] T. A. Bendikov, J. Kim, and T. C. Harmon, "Development and environmental application of a nitrate selective microsensor based on doped polypyrrole films," sensor. Actuat. B-Chem., vol. 106, no. 2, pp. 512–517, 2005.

[3] M. Sohail and S. B. Adeloju, "Nitrate biosensors and biological methods for nitrate determination," Talanta, vol. 153, pp. 83–98, 2016.

[4] X. F. Yue, Z. Q. Zhang, and H. T. Yan, "Flow injection catalytic spectrophotometric simultaneous determination of nitrite and nitrate," Talanta, vol. 62, no. 1, pp. 97–101, 2004.

[5] S. S. Chou, J. C. Chung, and D. F. Hwang, "A high performance liquid chromatography method for determining nitrate and nitrite levels in vegetables," J. Food Drug Anal., vol. 11, no. 3, pp. 233–238, 2003.

[6] M. P. N. Bui, J. Brockgreitens, S. Ahmed, and A. Abbas, "Dual detection of nitrate and mercury in water using disposable electrochemical sensors," Biosens. Bioelectron., vol. 85, pp. 280–286, 2016.

[7] L. Gao, J. Barber-Singh, S. Kottegoda, D. Wirtshafter, and S. A. Shippy, "Determination of nitrate and nitrite in rat brain perfusates by capillary electrophoresis," Electrophoresis, vol. 25, no. 9, pp. 1264–1269, 2004.

[8] Z. Legnerová, P. Solich, H. Sklenářová, D. Šatínský, and R. Karlíček, "Automated simultaneous monitoring of nitrate and nitrite in surface water by sequential injection analysis," Water Res., vol. 36, no. 11, pp. 2777–2783, 2002.

[9] C. Huber, I. Klimant, C. Krause, T. Werner, and O. S. Wolfbeis, "Nitrate-selective optical sensor applying a lipophilic fluorescent potential-sensitive dye," Anal. Chim. Acta, vol. 449, pp. 81–93, 2001.

[10] A. P. Turner, "Biosensors: sense and sensibility," Chem. Soc. Rev., vol. 42, no. 8, pp. 3184–3196, 2013.

[11] F. Can, S. K. Ozoner, P. Ergenekon, and E. Erhan, "Amperometric nitrate biosensor based on Carbon nanotube/Polypyrrole/Nitrate reductase biofilm electrode," Mater. Sci. Eng. C, vol. 32, no. 1, pp. 18–23, 2012.

[12] R. Ahmad, K. S. Bhat, M. S. Ahn, and Y. B. Hahn, "Fabrication of a robust and highly sensitive nitrate biosensor based on directly grown zinc oxide nanorods on a silver electrode," New J. Chem., vol. 41, no. 19, pp. 10992–10997, 2017.

[13] M. Al Mamun, Y. A. Wahab, M. M. Hossain, A. Hashem, A., and M. R. Johan, "Electrochemical biosensors with aptamer recognition layer for the diagnosis of pathogenic bacteria: barriers to commercialization and remediation," TrAC-Trend. Anal. Chem., p. 116458, 2021.

[14] P. Kalimuthu, K. Fischer-Schrader, G. Schwarz, and P. V. Bernhardt, "A sensitive and stable amperometric nitrate biosensor employing Arabidopsis thaliana nitrate reductase," J. Biol. Inorg. Chem., vol. 20, no. 2, pp. 385–393, 2015.

[15] F. Cui, Y. Yue, Y. Zhang, Z. Zhang, and H. S. Zhou, "Advancing biosensors with machine learning," ACS Sensors, vol. 5 no. 11, pp. 3346–3364, 2020.

[16] K. Asefpour Vakilian and J. Massah, "A portable nitrate biosensing device using electrochemistry and spectroscopy," IEEE Sens. J., vol. 18, no. 8, pp. 3080–3089, 2018.

[17] K. Asefpour Vakilian and J. Massah, "A fuzzy-based decision making software for enzymatic electrochemical nitrate biosensors," Chemometr. Intell. Lab., vol. 177, pp. 55–63, 2018.

[18] J. Massah and K. Asefpour Vakilian, "An intelligent portable biosensor for fast and accurate nitrate determination using cyclic voltammetry," Biosyst. Eng., vol. 177, pp. 49–58, 2019.

[19] I. Lambeck, J. C. Chi, S. Krizowski, S. Mueller, N. Mehlmer, M. Teige, K. Fischer, and G. Schwarz, "Kinetic analysis of 14-3-3-inhibited *Arabidopsis thaliana* nitrate reductase," Biochemistry, vol. 49, no. 37, pp. 8177–8186, 2010.

[20] P. Kalimuthu, K. Fischer-Schrader, G. Schwarz, and P. V. Bernhardt, "Mediated electrochemistry of nitrate reductase from *Arabidopsis thaliana*," J. Phys. Chem. B, vol. 117, no. 25, pp. 7569–7577, 2013.

[21] A. Asefpour Vakilian and M. R. Saradjian Maralan, "Optimizing the regularization parameters of prior information in sparse coding-based multispectral image fusion," Earth Observ. Geomat. Eng., vol. 4, no. 2, pp. 169–179, 2020.

[22] X. S. Yang, "Nature-inspired optimization algorithms: Challenges and open problems," J. Comput. Sci.-Neth., vol. 46, p. 101104, 2020.

[23] M. Khandelwal and D. J. Armaghani, "Prediction of drillability of rocks with strength properties using a hybrid GA-ANN technique," Geot. Geol. Eng., vol. 34, no. 2, pp. 605–620, 2016.

[24] V. Kumar, A. Kumar, D. Chhabra, and P. Shukla, "Improved biobleaching of mixed hardwood pulp and process optimization using novel GA-ANN and GA-ANFIS hybrid statistical tools," Bioresource Technol., vol. 271, pp. 274–282, 2019.

[25] G. F. Gomes, F. A. de Almeida, D. M. Junqueira, S. S. da Cunha Jr, and A. C. Ancelotti Jr, "Optimized damage identification in CFRP plates by reduced mode shapes and GA-ANN methods," Eng. Struct., vol. 181, pp. 111–123, 2019.

[26] P. Kalimuthu, P. Ringel, T. Kruse, and P. V. Bernhardt, "Direct electrochemistry of nitrate reductase from the fungus *Neurospora crassa*." BBA-Bioenergetics, vol. 1857, no. 9, pp. 1506–1513, 2016.

[27] D. A. Cataldo, M. Maroon, L. E. Schrader, and V. L. Youngs, "Rapid colorimetric determination of nitrate in plant tissue by nitration of salicylic acid," Commun. Soil Sci. Plan., vol. 6, pp. 71–80, 1975.

[28] K. Asefpour Vakilian, "Machine learning improves our knowledge about miRNA functions towards plant abiotic stresses," Sci. Rep.-UK, vol. 10, p. 3041, 2020.

[29] K. Asefpour Vakilian, "Gold nanoparticles-based biosensor can detect drought stress in tomato by ultrasensitive and specific determination of miRNAs," Plant Physiol. Bioch., vol. 145, pp. 195–204, 2019.

[30] K. Asefpour Vakilian, "Determination of nitrogen deficiency-related microRNAs in plants using fluorescence quenching of graphene oxide nanosheets," Mol. Cell. Probe, vol. 52, p. 101576, 2020.

[31] Z. Tabibi, J. Massah, and K. Asefpour Vakilian, "A biosensor for the sensitive and specific measurement of arsenite using gold nanoparticles," Measurement, vol. 187, p. 110281, 2022.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

ECG-Based Prediction of Epileptic Seizures Using Machine Learning Methods

1st Bahare Seifi[1], 2nd Mahsa Barfi[2], 3rd Mansour Esmaeilpour[3]

[1] *Computer Engineering Department, Hamedan Branch, Islamic Azad University, Hamedan, Iran,* seifi.bahar9@gmail.com

[2] *Electrical Engineering Department, Bu-Ali Sina University, Hamedan, Iran,* m.barfi@eng.basu.ac.ir

[3] *Computer Engineering Department, Hamedan Branch, Islamic Azad University,* Hamedan, Iran, esmaeilpour@iauh.ac.ir

***Abstract*— Epilepsy is a type of neurological disorder that is associated with recurrent seizures. This study aimed to present three machine learning methods for predicting epileptic seizures using the features of electrocardiogram (ECG) signals. First, the ECG data set consisting of 13 people were pre-processed. Then, 13 features were extracted. SVM, KNN, and Naive Bayes were used to classify the interictal and preictal sections of each patient's ECG signal. The prediction interval of 15 minutes was selected. The results show that the Naive Bayes is a suitable choice to predict epileptic seizures.**

***Index Terms* — Electrocardiogram, Epilepsy, K-nearest neighbors, Naive Bayes, Support vector machines, Wrapper method.**

I. INTRODUCTION

Epilepsy is a chronic disorder of the central nervous system that causes people to have recurrent seizures [1]. Five out of every 1,000 people suffer from epilepsy, the third most common neurological disorder in the United States after Alzheimer's disease and cerebrovascular accidents [2].

In some types of epilepsy, seizures are not controlled despite proper medication and treatment, called drug-resistant epilepsy. Therefore, predicting epileptic seizures is an essential issue to avoid possible risks [3].

Recent algorithms use electroencephalogram (EEG) and electrocardiogram (ECG) signals to detect the onset of a seizure. Three main categories of these algorithms are used to detect seizures: EEG-based seizure event detectors, EEG-based seizure initiation detectors, and EEG / ECG-based seizure detectors. Biotic Signal processing plays a vital role in the correct and timely prediction of epilepsy.

According to previous research, ECG signal analysis can help predict epileptic seizures. Some studies found significant changes along with the preictal parts of ECG signals by analyzing the parameters of heart rate variability and comparing them statistically in the preictal and interictal parts [4-6]. These changes indicate high sympathetic activity and occur especially five minutes before the onset of seizures [7]. Therefore, ECG signal analysis can be a beneficial tool for predicting epileptic seizures.

Epilepsy is often accompanied by changes in the autonomic nervous system that usually occur a few minutes before a seizure [8]. The aim of this study is to predict seizures in specific patients by examining the characteristics of ECG signals.

In this study, ECG signals from 15 people with epilepsy collected by the University of Siena were used. First, ECG signals are preprocessed. Some filters are applied to eliminate their noise and achieve signals with better quality. These filters include a notch filter to eliminate power line interference and a butter-worth filter to remove baseline wander. Then, by an algorithm, the R peaks of the ECG signals are detected, and the RRi series are obtained. In the next step, 13 features are extracted from these series. They include time-domain and frequency-domain features. Then the best features for classification are selected by wrapper method. In the last step, three classification methods including support vector machines (SVM), K-nearest neighbors (KNN), and Naive Bayes have been performed to classify the two parts of the preictal and interictal ECG signals, which are among the most widely used machine learning methods. The results of these three methods have been compared in terms of a number of statistical criteria. These criteria include correctly classified instances, TP rate, FP rate, precision, recall, and F-measure. These criteria for Naive Bayes are better than SVM and KNN, so Naive Bayes is a

suitable choice for predicting epileptic seizures. It should be noted that the Naive Bayes and SVM had almost the same performance.

Most studies use EEG signals to predict epilepsy, while this study uses ECG signals. Furthermore, an effective method called the wrapper method has been used to select the best features for classification. Finally, three widely used Learning machine methods have been used to classify and predict epileptic seizures.

II. DATA COLLECTION

Data were collected by the Department of Neurological and Neurosensorial Sciences, University of Siena, Italy, during a regional research project called PANACEE, aimed at developing low-cost patient-specific monitoring and control devices to predict epileptic seizures [7, 8].

The database includes EEG recordings from 15 patients monitored by Video-EEG with a sampling rate of 512 Hz. The electrodes were arranged on the basis of the international 10-20 system [2, 9]. Simultaneously, ECGs were recorded at a sampling frequency of 512 Hz.

The data set includes 15 patients. These 15 people (with 38 seizures) include eight women and seven men in the age range of 9.9 ± 17.6. Seizures are classified according to International League Against Epilepsy (ILAE) standards [10], while patients' resistance and response to treatment are determined according to the new ILAE 2010 [11] criteria [7, 8]. In this article, only ECG signals of these patients are used. A detailed description of the data set is provided in [7]. It should be noted that in this article, ECG data of 13 patients were analyzed. The data of Patients No. 1 and 2, because they had only one seizure and did not have good classification results, were not used.

The image of the seizure area of one of the patients is shown in Fig. 1.

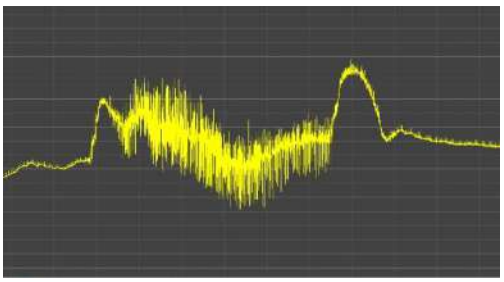


Figure 1. The seizure area of one of the patients.

III. ECG SIGNAL PROCESSING

The proposed method of the present study includes the pre-processing of ECG signals, the extraction of time-domain and frequency-domain features of the data, feature selection by wrapper method, and finally, the use of the KNN, SVM, and Naive Bayes for classification and performance evaluation. The flowchart of the proposed method is shown in Fig. 2.

A. Pre-processing

The frequency band of the ECG signal is between 0.15 and 150 Hz, but in recording, noise due to the muscle activity, disturbance of the baseline due to changes in the contact impedance of the electrode and the body, and offset potentials are also included. Baseline oscillations have a frequency of less than 1 Hz, and muscles noise has a frequency of more than 150 Hz [12, 13]. Therefore, filters on ECG signals were first used to eliminate powerline interference (50Hz) and remove baseline wander. A first-order low pass Butterworth filter with a cut-off frequency of 3.1 Hz was used to remove the baseline wander. A Notch filter with Resonator model, 1 Hz cut-off frequency, and Notch Q-factor 20 has been used to eliminate power line interference.

After the noise is removed, the signals are normalized using (1) to reduce interpersonal variability.

$$x_{Normalized\ Signal} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

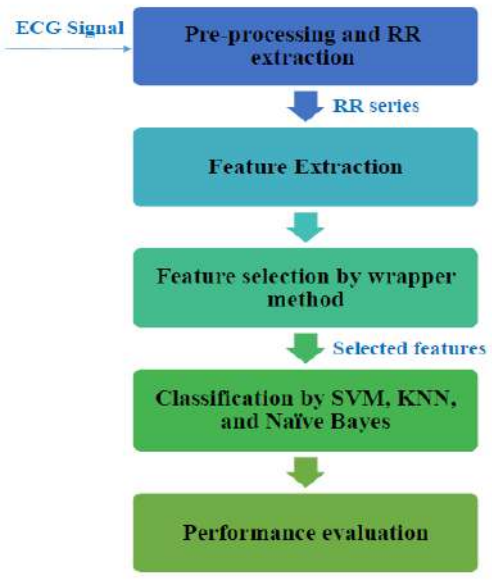


Figure 2. The flowchart of the proposed method.

B. Detection of R peaks and obtaining of RRi series

In this paper, SinusCor [14] software is used to detect R peaks, which is an advanced tool for analyzing heart rate variability. The software uses a simple derivative-based method to detect R-peaks. The detection process of these peaks is that, first, to eliminate the base noise and high frequencies, the ECG signal is filtered with a bandpass Butterworth filter with a default 5 to 200 Hz bandwidth. Then, if the first derivative of the filtered signal exceeds zero and the corresponding ECG point is greater than a fixed threshold, R peaks are detected [14].

During this process, the RRi series may be contaminated by artifacts or ectopic beats, thus affecting HRV analysis [15]. A widely used method to remove abnormal beats from the tachogram is to filter the RRi [15-17]. For this purpose, in this study, a moving-median filter with degree 3 is used.

After obtaining the RR series, these series are interpolated and then resampled at a sampling frequency of 4 Hz to arrange their sampling points at equal intervals.

C. Feature extraction

After pre-processing and detection of R peaks, RR series features were extracted for each ECG signal in 180-second windows with a 60-second overlap. According to recent research, short-term record analysis is a reliable source for evaluating the autonomic nervous system of the heart [18, 19]. A total of 13 properties were calculated, which include time-domain and frequency-domain features.

D. Time-domain features

- The time-domain features calculated in this research are as follows: [1, 7, 14, 20]:
- RMSSD: root mean square of successive differences that shows short-term signal changes.
- SDNN: standard deviation of RR intervals, which provides information about short-term and long-term signal changes.
- pNN50: The number of successive RRi pairs, whose difference is more than 50 ms, is divided into the total number of RRi intervals and multiplies by 100.
- SD1: Poincaré image standard deviation on the line perpendicular to the line of identity, which is a measure of short-term signal changes.
- SD2: Poincaré image standard deviation on the line of identity, which is a measure of long-term signal changes.
- RRi: RRi values.
- HRi: HRi values.

E. frequency-domain features

HRV analysis in frequency-domain are performed by spectral analysis of the RRi signal. First, the power spectrum density (PSD) is estimated using Welch's periodogram. The frequency domain criteria are then calculated.

In SinusCor software, it is possible to select PSD estimation through Welch periodogram [21]. In the Welch method, frequency components are calculated using the estimated mean PSD [21].

The frequency-domain features computed are the following [14]:

- VLF: Energy of the PSD from 0 to 0.04 Hz in absolute values.
- LF: Energy of the PSD from 0.04 to 0.15 Hz in absolute values.
- HF: Energy of the PSD from 0.15 to 0.4 Hz in absolute values
- LF/HF: Ratio of LF to HF. The LF/HF ratio is a crucial frequency-domain feature that determines the degree of sympathetic-parasympathetic balance of the autonomic nervous system (ANS).
- $LF_{n.u}$: LF energy normalized by total power and VLF.
- $HF_{n.u}$: HF energy normalized by total power and VLF.

F. Segmentation of feature vectors

After the feature extraction step, each feature vector was segmented in the following two states:

Preictal: 15 minutes before the onset of seizures until the onset of seizures.

Interictal: Non-seizure data before the preictal state, at least 50 minutes before the seizure [7, 22].

G. Feature selection

In this study, the wrapper method was used to select the ECG signal features of each patient. In this method, first, all possible subsets of features are created. A kind of classifier is then applied to these subsets of features. Each of the subsets of features that resulted in better classifying accuracy, that features subset is selected as the best features set for that classifier. This method is one of the most widely used methods in machine learning and low-dimensional data. Consequently, Therefore, in this study, due to the low dimensions of the data, the use of this method has led to appropriate classification results.

It is important to note that for the Naive Bayes classifier, the wrapper method performed better. But in both SVM and KNN, using all 13 features in most patients had better classification results than the wrapper method. The selected features for each patient and each classifier are shown in Table 1.

H. Classification

The purpose of this section is to classify ECG signals into two parts, preictal and interictal, based on the selected features.

In this research, three widely used approaches in machine learning including KNN, SVM, and Naive Bayes have been used. It should be remarked that these classification methods are used to predict epilepsy as patient-specific [7] methods

because the type of epilepsy of individuals and their seizure patterns are different.

- Support Vector Machines (SVM)

Support Vector Machines is a classifier that has a solid mathematical base and good performance with a small training data set [23]. The kernel function in a SVM classifier is usually a polynomial, Gaussian, RBF(radial basis function) or sigmoid function [24]. The SVM makes its predictions using a linear combination of the Kernel function that applies to a set of training data called support vectors [25, 26]. RBF function has been used as a SVM kernel for some patients, and in others, a polynomial function has been used. Any kind of kernel that has led to better classification results has been selected as the final kernel.

A 10-fold cross validation is used to classify new movements. In the cross-validation method, the data is divided into two groups: "training" and "test". The K-fold cross-validation randomly splits the data into K categories, each of which has the same number of samples [27]. Each category is selected once as test data. The classification results using SVM are shown in Table 2.

- K-Nearest Neighbors (KNN)

K-nearest neighbors [28] is one of the machine learning algorithms in which a sample is categorized by a majority vote of its neighbors. This sample is determined in the most general class among the k neighbors. K is a positive value and is generally small. Also, a 10-fold cross validation is used in this method. The classification results using KNN are shown in Table 3.

- Naive Bayes

Naive Bayes, proposed by the British scientist Thomas Bayes [29], is a classifier that uses probabilistic and statistical methods. The Naive Bayes simply allows all features to participate equally in the final decision, so it is a popular model in machine learning algorithms. Its advantages include simple computing, easy to implement, time efficiency, and not sensitive to data noise [30]. Also, a 10-fold cross validation is used in this method. The classification results using Naive Bayes are shown in Table 4.

- Performance evaluation

For each patient, the results of the three classification methods were evaluated based on the following criteria [31]:

- Correctly Classified Instance: The number of samples that are correctly classified, expressed as a percentage.
- TP Rate (True positive rate or Sensitivity): correctly classified interictal segments. This criterion is specified for each class separately.
- FP Rate (False positive rate) : incorrectly classified interictal segments. This criterion is specified for each class separately.
- Precision: this is a ratio between the true positives and all the positives.

Precision was calculated as:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

- Recall: The recall is the measure of a classifier correctly identifying True Positives.

Recall was calculated as:

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

Where FN is false negative rate.

- F-Measure: this is an appropriate criterion for assessing the accuracy of an experiment. F-measure is one in the best case and zero in the worst case.

F-measure was calculated as:

$$\text{F-measure} = \frac{2.recall.precision}{recall+precision} \quad (4)$$

IV. RESULTS AND DISCUSSION

In this study, we tried to address the main challenge of predicting seizures in epileptic patients. Because epilepsy is often associated with changes in the autonomic nervous system that usually begin a few minutes before the seizure onset, it is essential to identify these changes to avoid the risks and to have enough time before the seizure. In addition, these changes can be a risk factor for epilepsy patients and increase the risk of death. Seizure prediction is done on a patient-specific approach basis. This approach is more effective for predicting seizures because the physical condition and type of epilepsy of the patients are different. First, ECG data of 13 patients were pre-processed. Then, using an algorithm, R peaks are detected, and RRi series are obtained. Then 13 faetures are extracted from them, which include time-domain and frequency-domain properties. The next step is to select a feature by the wrapper method. Classification of two parts, Preictal, and interictal ECG signal, has been done by three algorithms: SVM, KNN, and Naive Bayes. The results of these three methods have been evaluated using correctly classified instances, TP rate, FP rate, precision, recall, and F-measure criteria. The prediction time (Preictal section) in this study was 15 minutes.

The average values of correctly classified instances (CCI), TP rate, FP rate, precision, recall, and F-measure for SVM classifier are 73.45, 73.45, 56.70, 71.81, 72.99, and 0.74 respectively. These criteria, which were also calculated using the KNN classifier, are as follows: 69.48, 69.48, 47.93, 70.08, 69.48, and 0.69. By Naïve Bayes classifier, they are 75.22, 75.22, 37.28, 75.58, 75.22, and 0.72.

The comparison of the SVM, KNN, and Naïve Bayes in terms of correctly classified instances (%), TP rate (%), and FP rate (%) is shown in Fig. 3 and terms of precision (%) and recall (%) is shown in Fig. 4.

From Fig. 3 and Fig. 4 and a comparison of the average values of the performance evaluation criteria of the three classifiers, the Naïve Bayes classifier performs better than the KNN and SVM. The feature selection method, the wrapper

method, also performed better for the Naive Bayes. The Naive Bayes classifier performed better than two other classifiers because it does not take data noise into account and all features participate equally in the final decision. These results indicate that seizures can be accurately predicted using the proposed classifier.

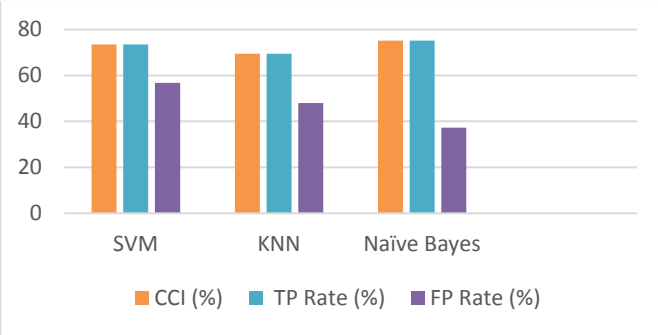


Figure 3. Comparison of SVM, KNN, and Naïve Bayes classifiers in terms of correctly classified instances (CCI), TP Rate, and FP Rate.

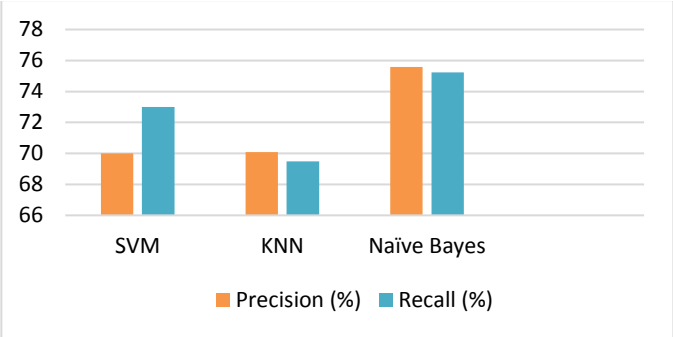


Figure 4. Comparison of SVM, KNN, and Naïve Bayes classifiers in terms of precision and recall.

TABLE 1: SELECTED FEATURES FOR EACH PATIENT AND EACH CLASSIFIER

Pat. No.	SVM	KNN	Naive Bayes
1	All of the 13 features	All of the 13 features	PNN50, RRi, and HRi
2	All of the 13 features	All of the 13 features	RMSSD, SDNN, and LF/HF
3	All of the 13 features	All of the 13 features	HRi
4	RMSSD and LF/HF	All of the 13 features	RMSSD and RRi
5	All of the 13 features	All of the 13 features	All of the 13 features
6	All of the 13 features	All of the 13 features	PNN50, RRi, and $LF_{n.u}$
7	All of the 13 features	All of the 13 features	RRi, LF, and $LF_{n.u}$
8	All of the 13 features	All of the 13 features	VLF
9	All of the 13 features	All of the 13 features	RMSSD, PNN50, and HRi
10	All of the 13 features	All of the 13 features	VLF
11	All of the 13 features	All of the 13 features	SDNN and SD1
12	All of the 13 features	All of the 13 features	RMSSD and LF/HF
13	RMSSD	All of the 13 features	RMSSD and HRi

TABLE 2: CLASSIFICATION RESULTS USING SVM

Pat. No.	Correctly Classified Instances (%)	TP Rate (Sensitivity) (%)	FP Rate (%)	Precision (%)	Recall (%)	F-Measure
1	67.56	67.60	58.90	63.30	67.60	0.61
2	68.18	68.20	74.40	51.90	68.20	0.59
3	85.71	85.70	10.70	89.30	85.70	0.85
4	81.81	81.80	48.50	85.50	81.80	0.78
5	50.00	50.00	64.30	31.80	50.00	0.38
6	71.64	71.60	71.60	71.60	71.60	0.83
7	77.41	77.40	77.40	77.40	77.40	0.87
8	71.73	71.70	64.60	79.90	71.70	0.61
9	69.56	69.60	69.60	69.60	69.60	0.82
10	94.11	94.10	19.10	94.50	94.10	0.93
11	69.56	69.60	69.60	69.60	69.60	0.82
12	69.56	69.60	69.60	69.60	69.60	0.82
13	78.04	78.00	38.90	79.60	78.00	0.75
Average	73.45	73.45	56.70	71.81	72.99	0.74

TABLE 3: CLASSIFICATION RESULTS USING KNN

Pat. No.	Correctly Classified Instances (%)	TP Rate (Sensitivity) (%)	FP Rate (%)	Precision (%)	Recall (%)	F-Measure
1	62.16	62.20	44.20	63.80	62.20	0.62
2	63.63	63.60	65.70	59.30	63.60	0.61
3	71.42	71.40	29.80	71.40	71.40	0.71
4	86.36	86.40	63.40	88.50	86.40	0.84
5	25.00	25.00	76.40	26.40	25.00	0.25
6	68.65	68.70	50.60	67.20	68.70	0.67
7	73.38	73.40	53.30	72.40	73.40	0.72
8	69.56	69.60	69.60	69.60	69.60	0.82
9	64.13	64.10	49.80	63.80	64.10	0.63
10	88.23	88.20	38.20	89.80	88.20	0.86
11	76.81	76.80	31.60	76.80	76.80	0.76
12	71.01	71.00	31.40	73.80	71.00	0.71
13	82.92	82.90	19.20	83.30	82.90	0.83
Average	69.48	69.48	47.93	70.08	69.48	0.69

TABLE 4: CLASSIFICATION RESULTS USING NAÏVE BAYES

Pat. No.	Correctly Classified Instances (%)	TP Rate (Sensitivity) (%)	FP Rate (%)	Precision (%)	Recall (%)	F-Measure
1	78.37	78.40	32.00	77.80	78.40	0.77
2	81.81	81.80	17.20	84.60	81.80	0.82
3	85.71	85.70	10.70	89.30	85.70	0.85
4	77.27	77.30	39.80	76.30	77.30	0.76
5	45.83	45.80	44.40	52.50	45.80	0.41
6	65.67	65.70	61.30	61.60	65.70	0.63
7	79.03	79.00	54.20	76.50	79.00	0.76
8	73.91	73.90	55.60	74.20	73.90	0.68
9	67.39	67.40	68.50	55.80	67.40	0.57
10	94.11	94.10	19.10	94.50	94.10	0.93
11	72.46	72.50	20.10	79.70	72.50	0.73
12	78.26	78.30	22.90	80.20	78.30	0.78
13	78.04	78.00	38.90	79.60	78.00	0.75
Average	75.22	75.22	37.28	75.58	75.22	0.72

REFERENCES

[1] S. Behbahani, N. J. Dabanloo, A. M. Nasrabadi, and A. Dourado, "Prediction of epileptic seizures based on heart rate variability," Technology and Health Care, vol. 24, no. 6, pp. 795-810, 2016.

[2] R. Abbasi, and M. Esmaeilpour, "Selecting Statistical Characteristics of Brain Signals to Detect Epileptic Seizures using Discrete Wavelet Transform and Perceptron Neural Network," International Journal of Interactive Multimedia & Artificial Intelligence, vol. 4, no. 5, 2017.

[3] A. Jacoby, and J. K. Austin, "Social stigma for adults and children with epilepsy," Epilepsia, vol. 48, pp. 6-9, 2007.

[4] A. Motie-Nasrabadi, S. Behbahani, N. J. Dabanloo, C. A. Teixeira, and A. Dourado, "Pre-ictal heart rate variability assessment of epileptic seizures by means of linear and non-linear analyses," Anadolu Kardiyoloji Dergisi-The anatolian journal of cardiology, 2013.

[5] E. Kolsal, et al., "Can heart rate variability in children with epilepsy be used to predict seizures?," Seizure, vol. 23, no. 5, pp. 357-362, 2014.

[6] M. Moridani, and H. Farhadi, "Heart rate variability as a biomarker for epilepsy seizure prediction," Clin. Study, vol. 3, no. 8, pp. 3-8, 2017.

[7] L. Billeci, D. Marino, L. Insana, G. Vatti, and M. Varanini. "Patient-specific seizure prediction based on heart rate variability and recurrence quantification analysis," PloS one, vol. 13, no. 9, pp. e0204339, 2018.

[8] L. Billeci, et al, "Epileptic seizures prediction based on the combination of EEG and ECG for the application in a wearable device," 2019 IEEE 23rd International Symposium on Consumer Technologies (ISCT), IEEE., 2019.

[9] S. Barra, A. Casanova, M. Fraschini, and M. Nappi, "EEG/ECG signal fusion aimed at biometric recognition," International conference on image analysis and processing, Springer, 2015.

[10] R. S. Fisher, et al., "Operational classification of seizure types by the International League Against Epilepsy: Position Paper of the ILAE Commission for Classification and Terminology," Epilepsia, vol. 58, no. 4, pp. 522-530, 2017.

[11] P. Kwan, et al., "Definition of drug resistant epilepsy: consensus proposal by the ad hoc Task Force of the ILAE Commission on Therapeutic Strategies," Wiley Online Library, 2010.

[12] M. Varanini, et al., “An efficient unsupervised fetal QRS complex detection from abdominal maternal ECG,” Physiological measurement, vol. 35, no. 8, pp. 1607, 2014.

[13] L. Billeci, and M. Varanini, “Characterizing electrocardiographic changes during pre-seizure periods through temporal and spectral features,” 2017 Computing in Cardiology (CinC), IEEE., 2017.

[14] R. Bartels, L. Neumamm, T. Peçanha, and A. R. S. Carvalho, “SinusCor: an advanced tool for heart rate variability analysis,” Biomedical engineering online, vol. 16, no. 1, pp. 1-15, 2017.

[15] N. Storck, M. Ericson, L. E. Lindblad, and M. Jensen-Urstad, “Automatic computerized analysis of heart rate variability with digital filtering of ectopic beats,” Clinical Physiology, vol. 21, no. 1, pp. 15-24, 2001.

[16] J. Piskorski, and P. Guzik, “Filtering poincare plots,” Computational methods in science and technology, vol. 11, no. 1, pp. 39-48, 2005.

[17] M. Karlsson, R. Hörnsten, A. Rydberg, and U. Wiklund, “Automatic filtering of outliers in RR intervals before analysis of heart rate variability in Holter recordings: a comparison with carefully edited data,” Biomedical engineering online, vol. 11, no. 1, pp. 1-12, 2012.

[18] H. M. Melo, et al., “Ultra-short heart rate variability recording reliability: The effect of controlled paced breathing,” Annals of noninvasive electrocardiology, vol. 23, no. 5, pp. e12565, 2018.

[19] R. L. Gomes, et al., “Poincaré plot analysis of ultra-short-term heart rate variability during recovery from exercise in physically active men,” The Journal of sports medicine and physical fitness, vol. 58, no. 7-8, pp. 998-1005, 2017.

[20] K. Fujiwara, et al., “Epileptic seizure prediction based on multivariate statistical process control of heart rate variability features,” IEEE Transactions on Biomedical Engineering, vol. 63, no. 6, pp. 1321-1332, 2015.

[21] P. Welch, “The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms,” IEEE Transactions on audio and electroacoustics, 1967. vol. 15, no. 2, pp. 70-73, 1967.

[22] M. Z. Parvez, and M. Paul, “Epileptic seizure prediction by exploiting spatiotemporal relationship of EEG signals using phase correlation,” IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 24, no. 1, pp. 158-168, 2015.

[23] Q. Huang, et al., “A novel unsupervised adaptive learning method for long-term electromyography (EMG) pattern recognition,” Sensors, vol. 17, no. 6, pp. 1370, 2017.

[24] J. M. Fajardo, J.M., O. Gomez, and F. Prieto, “EMG hand gesture classification using handcrafted and deep features,” Biomedical Signal Processing and Control, vol. 63, pp. 102210, 2021.

[25] P. Detti, et al., “A patient-specific approach for short-term epileptic seizures prediction through the analysis of EEG synchronization,” IEEE Transactions on Biomedical Engineering, vol. 66, no. 6, pp. 1494-1504, 2018.

[26] M. Qaraqe, M. Ismail, E. Serpedin, and H. Zulfi, “Epileptic seizure onset detection based on EEG and ECG data fusion,” Epilepsy & Behavior, vol. 58, pp. 48-60, 2016.

[27] M. H. Bollen, and I.Y. Gu, “Signal processing of power quality disturbances,” John Wiley & Sons, vol. 30, 2006.

[28] Y. Narayan, “SEMG signal classification using KNN classifier with FD and TFD features,” Materials Today: Proceedings, vol. 37: pp. 3219-3225, 2021.

[29] H. Muhamad, C. A. Prasojo, N. A. Sugianto, L. Surtiningsih, and I. Cholissodin, “Optimasi Naïve Bayes Classifier Dengan Menggunakan Particle Swarm Optimization Pada Data Iris,” Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK), p-ISSN, pp. 2355-7699, 2017.

[30] A. P. Wibawa, et al., “Naïve Bayes Classifier for Journal Quartile Classification,” Int. J. Recent Contributions Eng. Sci. IT, vol. 7, no. 2, pp. 91-99, 2019.

[31] T. Thilagaraj, and N. Sengottaiyan, “Implementation of Classification Algorithms in Educational Data using Weka Tool,” 2019.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Fractional entropy and its applications in fuzzy c-means clustering

R. Beigmohamadi

Department of Mathematics, Institute for Advanced Studies in Basic Sciences, Zanjan 45137-66731, Iran,
rbegmohamadi@gmail.com

***Abstract*— In this work, a novel fuzzy clustering method based on fuzzy c-means clustering is introduced. Indeed, in order to maximize the dissimilarity between clusters a fractional entropy, as a regularization function, is added to its objective function. Then, an explicit expression to update the partition matrix is presented.**

***Index Terms*— Fractional entropy, Fuzzy clustering, Fractional entropy fuzzy clustering.**

I. Introduction

In order to find regularities in observation pattern recognition, which is a compilation of computer procedure, is commonly used. There exist varied approaches for this technique, clustering or unsupervised classification and supervised classification. The natural grouping exists in observations can be obtained using clustering methods. This is implemented in a way that objects belonging to the different clusters are as dissimilar as possible and objects belonging to the same cluster are as similar as possible. A cluster is usually represented as either a prototype data instance nearest to the center (called centroid) or grouping of similar data points around centroid. In other words, a cluster can be considered either with or without a well-defined boundary. Clusters without well-defined boundaries are called fuzzy clusters, while those with such feature are called crisp clusters.

Imprecision and uncertainties in real world phenomena may be handled using fuzzy logic that was introduced by Zadeh. In fact, two human capabilities can be formalized/mechanized by taking into account of fuzzy logic theory. The first one is the capability of reasoning and making logical decisions in an environment of imprecision. the second is the capability of performing broad diversity of physical and mental assignments without any computations and any measurements [24]. Clustering methods using fuzzy logic make the possibility of belonging each observation to more than one cluster. This combination would provide an appropriate tools to consider the uncertainties in real world observations.

One of the most commonly used fuzzy clustering schemes is Fuzzy C-means (FCM) algorithm, which was originally introduced by Dunn [7] and had been modified by Bezdek [2]. In FCM, the membership degrees of a data point are iteratively updated in which the number of clusters determined in advanced. Thus, all clusters possess a data point with the corresponding membership degree.

During the last decades the generalization of concepts such as entropy and information has been attracted the attention of scientific communities [18]. Clausius and Boltzmann introduced the concept of entropy in thermodynamics, and then, Shannon and Jaynes adopted entropy in information theory [1].

Recently, fuzzy clusters based on entropy are constructed by means of a similarity-threshold value, which is a very popular technique to increase the similarity between clusters [3], [26].

The distance between two distributions may be measured by relative entropy that is a general form of entropy. Many researcher have applied this function as a regularization function in FCM such as, Miyagishi et al. [17], Ichihashi et al. [11], and Ichihashi et al. [12]. The new objective function that contains entropy and relative entropy has remarkable advantageous such as controlling the fuzziness of the output clusters, assigning more consistent notions for semi-supervised algorithm, verifying the kernel weights for unsupervised algorithm [19].

Since the late 19th century, many researcher have paid attention to the development of fractional calculus. Non-local structure of fractional operators deals with memory effects in mathematical models. Accordingly, the theory of fractional calculus is an effective mathematical tool to model complex dynamic systems which involve memory effects. Recently, researchers brought together fractional calculus and information theory and proposed some generalized concepts of entropy [8], [16], [23].

In this work, we aim at introducing a new fuzzy clustering method based on fractional entropy. Indeed, we add a novel fractional entropy introduced in [16], as a regularization function in FCM objective function. Then, we obtain an explicit formula for updating the partition matrix.

II. A GENERALIZATION OF ENTROPY BASED ON FRACTIONAL CALCULUS

The development of information theory had been provided by Claude Shannon in 1948 [20], [21], and this theory has been used in many scientific fields. The information content of some events with probability q_i is considered as the basic cornerstone:

$$I(q_i) = -\ln(q_i). \tag{1}$$

Therefore, Shannon entropy, which is the expected value, have the following representation

$$S = E(-\ln(q)) = \sum_i (-\ln(q_i))q_i, \tag{2}$$

where $E(\cdot)$ denotes the expected value operator.

The Shannon entropy (2) conforms the four Khinchin axioms [13], [14], while only a sub-set of Khinchin axioms was conformed by several generalizations of entropy.

Recently, taking into account of the fractional calculus, Machado [16] proposed a general expression for Shannon entropy, which is defined by:

$$S_\nu = \sum_i q_i \left\{ -\frac{q_i^{-\nu}}{\Gamma(\nu+1)} \left[\psi(1) - \psi(1-\nu) + \ln(q_i)\right] \right\}, \tag{3}$$

where $\psi(\cdot)$ and $\Gamma(\cdot)$ represent the digamma and gamma functions, and ν is the fractional derivative order.

Moreover, the fractional order information can be determined by [16]

$$I_\nu = -\frac{q_i^{-\nu}}{\Gamma(\nu+1)} \left[\psi(1) - \psi(1-\nu) + \ln(q_i)\right]. \tag{4}$$

Fractional entropy may be viewed as an innovative formula for entropy that inherited its characterization from the theory of fractional calculus.

Some of the Khinchin axioms are not conformed by the fractional entropy (3), except for the case $\nu = 0$, which implies the classical Shannon entropy. This behavior have relevance with the properties of fractional derivative operator, since fractional derivative operator does not possess some properties of the integer order derivative. In other word, we loose some classical properties by generalizing the operators in both cases. For more detail, we refer the reader to see [16].

III. FRACTIONAL FUZZY ENTROPY CLUSTERING METHOD

Entropy could be used as an appropriate tool to measure the uncertainty and impression appeared in real world phenomena. In fact, entropy is the minimum descriptive complexity of a random variable [5].

On the basis of the well-known FCM we design a new clustering method in which the additional term in its objective function is the fractional order entropy (4). This new model has some advantageous. First, the probabilistic interpretation of membership functions in FCM caused by the probabilistic constraint, in which the summation of membership degrees for an observation in all clusters must be 1 [6].

Entropy and relative entropy as the basic quantities of information theory are also defined as a function of probability distributions [5], and therefore, the theory of entropy and FCM share the same notions. Second, the convexity and non-negativity properties of entropy make it a proper candidate to be used as regularization function. Moreover, according to [10], the complexity of fractional fuzzy entropy clustering algorithm is more accurate, especially those time series with high complexity. Therefore, we propose a new fractional order entropy function as a regularization function in FCM in which the effect of other clusters are not ignored. Based on this idea, we design the proposed clustering method.

Consider u_{ij} as the membership degree of jth observation in ith cluster. We define the following optimization model with the optimum value of u_{ij} as the desired solution.

$$\begin{aligned} \min JF(U,V,c) &= \sum_{i=1}^{c}\sum_{j=1}^{N} u_{ij}^m d_{ij}^2 \\ &+ \theta \left(\sum_{i=1}^{c}\sum_{j=1}^{N} u_{ij} \left[\frac{u_{ij}^{-\nu}}{\Gamma(\nu+1)} \left[\psi(1) - \psi(1-\nu) + \ln(u_{ij})\right] \right] \right), \end{aligned} \tag{5}$$

where c is the number of cluster, d_{ij}^2 is the distance of jth observation to ith cluster, m is the degree of fuzziness, N is the number of observations, ν is the fractional derivative order, and θ is positive number.

The distance between the cluster centers and observations should be as small as possible and this is the expected results in the first term in (5), whereas the second term maximize the entropy of the membership degree of ith clusters.

Analogous to FCM, we have the following constraint for u_{ij} .

$$\begin{cases} \sum_{i=1}^{c} u_{ij} = 1, & \forall j, \\ 0 < \sum_{j=1}^{N} u_{ij}, & \forall i, \\ u_{ij} \in [0,1], & \forall i,j. \end{cases} \tag{6}$$

It was indicated in some studies that one can use the entropy coefficient of membership degrees instead of fuzzifier m, because there no precise meaning and physical interpretation [22]. Therefore, we use the entropy coefficient because the fuzzy rate of membership degree has a close physical meaning with it [22]. In order to obtain an explicit formula to update the partition matrix it is required to consider the fuzzifier equal to 1. Otherwise, for any other values of fuzzifier one may lead to use numerical schemes to obtain the updated partition matrix.

Theorem 1. *The solution to the optimization problem* (5) *with the constraints* (6) *is given by*

$$u_{ij} = \left(\frac{-\frac{\theta(\nu-1)}{\Gamma(\nu+1)} W\left(-\frac{\nu(d_{ij}^2-\lambda_j)\left(e^{(\nu-1)(\psi(1)-\psi(1-\nu)-1)}\right)^{-\frac{\nu}{\nu-1}}}{\theta(\nu-1)} \right)}{\nu(d_{ij}^2-\lambda_j)} \right)^{\frac{1}{\nu}} \tag{7}$$

Moreover, the center points are updated by:

$$v_i = \frac{\sum_{j=1}^{N} u_{ij} x_j}{\sum_{j=1}^{N} u_{ij}}, \tag{8}$$

where $W(\cdot)$ is the Lambert-W function, $\lambda_j, j = 1,2,\ldots,N$ is the Lagrangian multiplier, x_j is the jth observation.

Proof. Minimizing problem

subject to the constraints (6) is a nonlinear optimization problem. A commonly used method to solve such problems is the Lagrangian multiplier technique. Therefore, we add the constraint term $(\sum_{i=1}^{c} u_{ij} = 1, \forall j)$ in objective function in (5) $(\lambda_j, j = 1, 2, \ldots, N)$, and obtain

$$\begin{aligned}\min JF' = &\sum_{i=1}^{c}\sum_{j=1}^{N} u_{ij} d_{ij}^2 \\ &+ \theta\left(\sum_{i=1}^{c}\sum_{j=1}^{N} u_{ij}\left[\frac{u_{ij}^{-\nu}}{\Gamma(\nu+1)}\left[\psi(1)-\psi(1-\nu)+\ln(u_{ij})\right]\right]\right) \\ &- \sum_{j=1}^{N}\lambda_j\left(\sum_{i=1}^{c} u_{ij} - 1\right). \end{aligned} \tag{9}$$

Since there is no correlation between the membership degrees of the observations [15], thus, minimizing (9) subject to U is equivalent to the problem in which one minimize the individual objective function regarding each u_{ij}, namely,

$$\begin{aligned} L = &u_{ij} d_{ij}^2 \\ &+ \theta\left(u_{ij}\left[\frac{u_{ij}^{-\nu}}{\Gamma(\nu+1)}\left[\psi(1)-\psi(1-\nu)+\ln(u_{ij})\right]\right]\right) \\ &- \lambda_j(u_{ij} - 1). \end{aligned} \tag{10}$$

One of the required conditions that optimizes the objective function (5) is $\frac{\partial J}{\partial U} = 0$ or equivalently $\frac{\partial L}{\partial u_{ij}} = 0$, so,

$$\begin{aligned} d_{ij}^2 + \frac{\theta u_{ij}^{-\nu}}{\Gamma(\nu+1)}\left((1-\nu)\left(\psi(1)-\psi(1-\nu)+\ln(u_{ij})\right)+1\right) \\ - \lambda_j = 0. \end{aligned} \tag{11}$$

Solving (11) with respect to u_{ij} would result in:

$$u_{ij} = \left(\frac{\theta(1-\nu)W\left(\frac{\nu(d_{ij}^2-\lambda_j)\left(e^{(\nu-1)(\psi(1)-\psi(1-\nu)-1)}\right)^{-\frac{\nu}{\nu-1}}}{\theta(1-\nu)}\right)}{\Gamma(\nu+1)\nu(d_{ij}^2-\lambda_j)}\right)^{\frac{1}{\nu}} \tag{12}$$

where $W(\cdot)$ is defined to be the function satisfying [4]:

$$W(Z)e^{W(Z)} = Z,$$

and called Lambert-W function. The second required condition that optimizes the objective function for is $\frac{\partial J}{\partial \lambda_j} = 0$, so:

$$\sum_{i=1}^{c}\left(\frac{\theta(1-\nu)W\left(\frac{\nu(d_{ij}^2-\lambda_j)\left(e^{(\nu-1)(\psi(1)-\psi(1-\nu)-1)}\right)^{-\frac{\nu}{\nu-1}}}{\theta(1-\nu)}\right)}{\Gamma(\nu+1)\nu(d_{ij}^2-\lambda_j)}\right)^{\frac{1}{\nu}} = 1. \tag{13}$$

Since the Lagrangian multiplier is totally related to the membership degree, it is not possible to solve the equation explicitly with respect to λ_j. Therefore, we aim at finding bounds for λ_j. There are two point of view to study this problem, $(1) d_{ij}^2 - \lambda_j > 0\ \forall i, j$ and $(2) d_{ij}^2 - \lambda_j < 0\ \forall i, j$. $(1) u_{ij} \geq 0\ \forall i, j$.

If $d_{ij}^2 - \lambda_j > 0$, then, $W(\cdot)$ must be non-negative, i.e.,

$$W\left(\frac{\nu(d_{ij}^2-\lambda_j)\left(e^{(\nu-1)(\psi(1)-\psi(1-\nu)-1)}\right)^{-\frac{\nu}{\nu-1}}}{\theta(1-\nu)}\right) \geq 0.$$

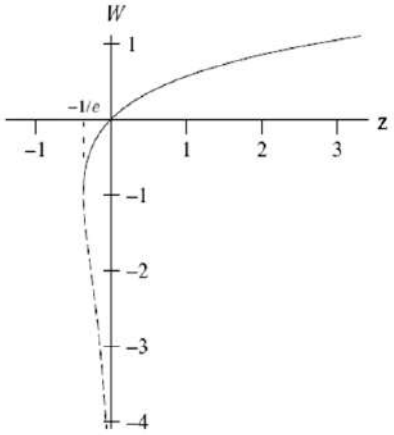


Fig. 1. Lambert-W function:—.

So, according to Fig. 1, we have,

$$\frac{\nu(d_{ij}^2-\lambda_j)\left(e^{(\nu-1)(\psi(1)-\psi(1-\nu)-1)}\right)^{-\frac{\nu}{\nu-1}}}{\theta(1-\nu)} \geq 0. \tag{14}$$

Since $d_{ij}^2 - \lambda_j > 0$ implies (14), so $u_{ij} \geq 0$.

Moreover, we must show that $u_{ij} \leq 1$. So, we have

$$\frac{\theta(1-\nu)W\left(\frac{\nu(d_{ij}^2-\lambda_j)\left(e^{(\nu-1)(\psi(1)-\psi(1-\nu)-1)}\right)^{-\frac{\nu}{\nu-1}}}{\theta(1-\nu)}\right)}{\Gamma(\nu+1)\nu(d_{ij}^2-\lambda_j)} \leq 1. \tag{15}$$

Then,

$$\begin{aligned} W\left(\frac{\nu(d_{ij}^2-\lambda_j)\left(e^{(\nu-1)(\psi(1)-\psi(1-\nu)-1)}\right)^{-\frac{\nu}{1-\nu}}}{\theta(1-\nu)}\right) \\ \leq \frac{\nu(d_{ij}^2-\lambda_j)\Gamma(\nu+1)}{\theta(1-\nu)}, \end{aligned} \tag{16}$$

According to the characteristic of the Lambert-W function the inequality (16) holds.

$(2) d_{ij}^2 - \lambda_j < 0$: then, $W(\cdot)$ must be non-positive, i.e.,

$$W\left(\frac{\nu(d_{ij}^2-\lambda_j)\left(e^{(\nu-1)(\psi(1)-\psi(1-\nu)-1)}\right)^{-\frac{\nu}{\nu-1}}}{\theta(1-\nu)}\right) \leq 0.$$

The characteristic of Lambert-W function imply that

$$-\frac{1}{e} < \frac{\nu(d_{ij}^2-\lambda_j)\left(e^{(\nu-1)(\psi(1)-\psi(1-\nu)-1)}\right)^{-\frac{\nu}{\nu-1}}}{\theta(1-\nu)} < 0. \tag{17}$$

Therefore, we get

$$d_{ij}^2 < \lambda_{ij} < \frac{\theta(1-\nu)}{e\nu}\left(e^{(\nu-1)(\psi(1)-\psi(1-\nu)-1)}\right)^{\frac{\nu}{\nu-1}} + d_{ij}^2. \tag{18}$$

As each individual observation must have a membership degree less than or equal to 1 in each cluster, so we have

$$\frac{\theta(1-\nu)W\left(\frac{\nu(d_{ij}^2-\lambda_j)\left(e^{(\nu-1)(\psi(1)-\psi(1-\nu)-1)}\right)^{-\frac{\nu}{\nu-1}}}{\theta(1-\nu)}\right)}{\Gamma(\nu+1)\nu(d_{ij}^2-\lambda_j)} \leq 1. \tag{19}$$

Thus, we have

$$W\left(\frac{\nu(d_{ij}^2-\lambda_j)\left(e^{(\nu-1)(\psi(1)-\psi(1-\nu)-1)}\right)^{-\frac{\nu}{\nu-1}}}{\theta(1-\nu)}\right) \geq \frac{\nu(d_{ij}^2-\lambda_j)\Gamma(\nu+1)}{\theta(1-\nu)}. \quad (20)$$

So, a direct calculation implies that another lower bound for λ_j is

$$\lambda_j > d_{ij}^2 + \frac{\theta(1-\nu)}{\nu\Gamma(\nu+1)}. \quad (21)$$

Therefore, based on (18) and (21), the lower and upper bound for λ_j would be:

$$d_{ij}^2 + \frac{\theta(1-\nu)}{\nu\Gamma(\nu+1)} < \lambda_j < \frac{\theta(1-\nu)}{e\nu}\left(e^{(\nu-1)(\psi(1)-\psi(1-\nu)-1)}\right)^{\frac{\nu}{\nu-1}} + d_{ij}^2. \quad (22)$$

The center point of each cluster is another main part of any clustering methods that needs to be optimized and updated. To do that, one can differentiate (10) w.r.t the center points, v_i, i.e. $\frac{\partial L}{\partial v_i}$. In this work, the Euclidean distance function are considered in d_{ij}^2, so the following formula updates the center points.

$$v_i = \frac{\sum_{j=1}^{N} u_{ij}x_j}{\sum_{j=1}^{N} u_{ij}}, \quad (23)$$

where x_j is the jth observation. □

There are two aspect to evaluate the implementation of the proposed method. The first one is to evaluate the computational complexity of the proposed method. In the second place, we will analyze the implementation of the new clustering method with various situations.

REFERENCES

[1] A. Ben-Naim, Statistical Thermodynamics Based on Information: A Farewell to Entropy; World Scientific: Singapore, Singapore, 2008.

[2] J.C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms. Kluwer Academic Publishers, Norwell, MA, USA, 1981.

[3] S. Chattopadhyay, D.K. Pratihar, S. Sarkar, A comparative study of fuzzy c-means algorithm and entropybased fuzzy clustering algorithms, Computing and Informatics, 30 (2011) 1001–1020.

[4] R.M. Corless, G.H. Gonnet, D.E.G. Hare, D.J. Jeffrey, D.E. Knuth, On the LambertW function, Advances in Computational Mathematics, 5 (1996) 329–359.

[5] T.M. Cover, J. Thomas, Elements of Information, John Wiley & Sons, New Jersey, 2006.

[6] J.V. De Oliveira, W. Pedrycz, Advances in Fuzzy Clustering and its Applications, John Wiley & Sons, 2007.

[7] J.C. Dunn, A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters. J. Cybernet, 3 (1973) 32—57.

[8] R.A.C. Ferreira, J.T. Machado, An Entropy Formulation Based on the Generalized Liouville Fractional Derivative, Entropy, 21 (2019).

[9] R.M. Gray, Entropy and Information Theory, Springer Verlag, 2010.

[10] S. He, K. Sun, R. Wang, Fractional fuzzy entropy algorithm and the complexity analysis for nonlinear time series, Eur. Phys. J. Special Topics, 227 (2018) 943–957

[11] H. Ichihashi, K. Honda, N. Tani, Gaussian mixture PDF approximation and fuzzy c-means clustering with entropy regularization, in: 4th Asian Fuzzy System Symposium,(2000) 217–221.

[12] H. Ichihashi, K. Miyagishi, K. Honda, Fuzzy c-means clustering with regularization by KL information, in: The 10th IEEE International Conference on Fuzzy Systems, 3 (2001) 924—927.

[13] E.T. Jaynes, Information theory and statistical mechanics. Phys. Rev. 106 (1957) 620–630.

[14] A.I. Khinchin, Mathematical Foundations of Information Theory; Dover: New York, NY, USA, 1957.

[15] R. Krishnapuram, J.M. Keller, A possibilistic approach to clustering, IEEE Transactions on Fuzzy Systems, 1 (1993) 98–110.

[16] J.T. Machado, Fractional Order Generalized Information, 16 (2014) 2350–2361.

[17] K. Miyagishi, H. Ichihashi, K. Honda, Fuzzy c-means clustering with regularization by KL information, Journal of Japan Society for Fuzzy Theory and Systems, 13 (2001) 406—417.

[18] A. Plastino, A.R. Plastino, Tsallis entropy and Jaynes' Information Theory formalism. Braz. J. Phys. 29 (1999) 50–60.

[19] F. Salehi, M.R. Keyvanpour, A. Sharifi, SMKFC-ER: Semi-supervised multiple kernel fuzzy clustering based on entropy and relative entropy, Information Sciences, 547 (2021) 667–688

[20] C.E. Shannon, A mathematical theory of communication. Bell Syst. Tech. J. 27 (1948) 379–423.

[21] C.E. Shannon, A mathematical theory of communication. Bell Syst. Tech. J. 27 (1948) 623–656.

[22] D. Tran, M. Wagner, Fuzzy entropy clustering, Ninth IEEE International Conference on Fuzzy Systems. FUZZ-IEEE 2000 (Cat. No. 00CH37063), vol 1, IEEE, (2000) 152–157.

[23] M.R. Ubriaco, Entropies based on fractional calculus. Phys. Lett. A 373 (2009) 2516–2519.

[24] L.A. Zadeh, Is there a need for fuzzy logic?, Information Sciences 178 (2008) 2751—2779

[25] M. Zarinbal, M.H. Fazel Zarndi, I.B. Turksen, Relative entropy fuzzy c-means clustering, Information Sciences, 260 (2014) 74–97.

[26] J. Yao, M. Dash, S.T. Tan, H. Liu, Entropy-Based Fuzzy Clustering and Fuzzy Modeling. Fuzzy Sets and System. 113 (2000) 381–388.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Fuzzy stability of involutions via fixed point technique

1st Parvaneh Lo′lo′, 2nd Ehsan Movahednia

1st Department of Mathematics, Behbahan Khatam Alanbia University of Technology, Behbahan, Iran, lolo@bkatu.ac.ir
2nd Department of Mathematics, Behbahan Khatam Alanbia University of Technology, Behbahan, Iran, movahednia@bkatu.ac.ir

***Abstract*— In this paper, we estabilish the generalized Hyers-Ulam-Rassias stability of involutions on fuzzy Banach algebras. Moreover, we show that under some conditions on an approximate involution, the fuzzy Banach algebra has a fuzzy C^*-algebra structure.**

***Index Terms*—fuzzy Banach $\star$-algebra, fuzzy C^*-algebra, Hyers-Ulam-Rassias stability**

I. Introduction

The stability question was first raised by Ulam [6], about the stability of group homomorphisms. This query has been positively responded for Banach space by Hyers [2]. He has asserted the following consequence. Let X_1 and X_2 be Banach space and let $k : X_1 \to X_2$ be mapping satisfying:

$$||k(v+w) - k(v) - k(w)|| \leq \varepsilon,$$

for some $\varepsilon > 0$ and for all $v, w \in X_1$. Then, there exists a unique additive mapping $H : X_1 \to X_2$ with:

$$||k(v) - H(v)|| \leq \varepsilon$$

for all $v \in X_1$. Hyers proved that

$$H(v) = \lim_{n\to\infty} 2^{-n} k(2^n x).$$

In 1987, Rassias [5] was the first to affirm the stability of the linear mapping between Banach spaces subject to a continuity assumption on the mapping. Recently, several fuzzy stability problems of various functional equations and their generalizations have been studied by some mathematicians.
In this paper, we estabilish the generalized Hyers-Ulam-Rassias stability of involutions on fuzzy Banach algebras. Further, we show that under some provisions on an approximate involution, the fuzzy Banach algebra has a fuzzy C^*-algebra structure.

Definition I.1. *Let X be a set. A function $d : X^2 \to [0, \infty]$ is a named a generalized metric on X if and only if d satisfies*

(G_1) $d(v, w) = 0 \Longleftrightarrow v = w \quad \forall v, w \in X$;
(G_2) $d(v, w) = d(w, v) \quad \forall v, w \in X$;
(G_3) $d(v, w) \leq d(v, u) + d(u, w) \quad \forall u, v, w \in X$.

Definition I.2. *Let X be areal vactor space. A function $N : X \times \mathbb{R} \to [0, 1]$ is named a fuzzy norm on X if for all $v, w \in X$ and all $p, q \in \mathbb{R}$*

(F_1) $N(v, p) = 0$, *for all* $p \leq 0$;
(F_2) $v = 0$ *if and only if* $N(v, p) = 1$ *for all* $p > 0$;
(F_3) $N(cv, p) = N(v, \frac{p}{|c|})$ *if* $c \neq 0$;
(F_4) $N(v + w, p + q) \geq \min\{N(v, p), N(w, q)\}$;
(F_5) $N(v, .)$ *is a non-decreasing function of* $\mathbb{R}$ *and* $\lim_{p\to\infty} N(v, p) = 1$;
(F_6) *For* $v \neq 0$, $N(v, .)$ *is continuous on* $\mathbb{R}$.

The pair (X, N) is named a fuzzy normed vector space.

Example I.1. *[3] Let $(X, ||.||)$ be a normed linear space and $\alpha, \beta > 0$. Define $N_1 : X \times \mathbb{R} \to [0, 1]$ by*

$$N_1(v, p) = \begin{cases} \frac{\alpha p}{\alpha p + \beta ||v||} & p > 0, v \in X, \\ 0 & p \leq 0, v \in X. \end{cases} \tag{1}$$

It is easy to check that N_1 is fuzzy norm on X.

Example I.2. *[3] Let $(X, ||.||)$ be a normed linear space and $\alpha > 0$. Define $N_2 : X \times \mathbb{R} \to [0, 1]$ by*

$$N_2(v, p) = \begin{cases} \frac{p^\alpha}{p^\alpha + ||v||^\alpha} & p > 0 \\ 0 & t \leq 0. \end{cases} \tag{2}$$

It is easy to check that N_2 is fuzzy norm on X.

Definition I.3. *Let (X, N) be a fuzzy normed vector space.*

(1) *A sequence $\{v_n\}$ in X is said to be convergent if there exists an $v \in X$ such that*

$$\lim_{n\to\infty} N(v_n - v, p) = 1, \quad \forall p > 0.$$

In this case, v is named the limit of the sequence $\{v_n\}$ and we denote it by $N - \lim_{n\to\infty} v_n = v$.

(2) *A sequence $\{v_n\}$ in X is named Cauchy if for each $\varepsilon > 0$ and each $p > 0$ there exists an $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ and all $m > 0$, we have*

$$N(v_{n+m} - v_n, p) > 1 - \varepsilon.$$

It is known that every convergent sequence in fuzzy normed space is Cauchy. If each Cauchy sequence is convergent, then the fuzzy norm is said to be complete and the fuzzy normed space is named a fuzzy Banach space. We say that a mapping $k : V \to W$ between fuzzy normed vactor spaces V and W is continuous at a point $v_0 \in X$ if for each sequence $\{v_n\}$ converging to v_0 in V, then the sequence $\{f(v_n)\}$ converges to $f(v_0)$. If $k : V \to W$ is continuous at each $v_0 \in V$, then $k : V \to W$ is said to be continuous on V.

Definition I.4. *[4] Let A be an algebra and (A, N) a fuzzy normed space. The fuzzy normed space (A, N) is called a fuzzy normed algebra if*

$$N(aa', st) \geq N(a, s).N(a', t) \qquad \forall a, a' \in A, \forall s, t \in \mathbb{R}^+.$$

Complete fuzzy normed algebra is called a fuzzy Banach algebra.

Example I.3. *Every normed algebra $(A, ||.||)$ defines a fuzzzy normed algebra (A, N), where N is defined by*

$$N(a', s) = \frac{s}{s + ||a'||} \qquad \forall a' \in A, \forall s > 0.$$

This space is called the induced fuzzy normed algebra.

Theorem I.1. *Let (X, d) be a generalized complete metric space. Assume that $G : X \to X$ is a strictly contractive operator with the Lipschitz constant $L < 1$. If there exists a nonnegative integer n_0 such that $d(G^{n_0+1}x, G^{n_0}x) < \infty$ for some $x \in X$, then the following statments are correct:*

(i) The sequence $\{G^n x\}$converges to a fixed point x_0 of G;
(ii) x_0 is the unique fixed point of G in $Y = \{y \in X | d(G^{n_0}x, y) < \infty\}$;
(iii) If $y \in Y$, then

$$d(y, x_0) \leq \frac{1}{1 - L} d(Gy, y).$$

II. Main result

We start the section with the following imprtant definitions.

Definition II.1. *Let A be a an algebra over $\mathbb{C}$. An involution on A is a mapping*

$$\begin{aligned} \star : & A \to A \\ & a \longmapsto a^\star \end{aligned}$$

such that

(i) $(\alpha a + \beta b)^\star; = \bar{\alpha} a^\star + \bar{\beta} b^\star \qquad \forall a, b \in A, \forall \alpha, \beta \in \mathbb{C}$;
(ii) $(ab)^\star = b^\star a^\star \qquad \forall a, b \in A$;
(iii) $a^{\star\star} = a \qquad \forall a \in A$.

1. A complex algebra with an involution is a $\star$-algebra.
2. A $C^\star$-algebras is a (non-zero) Banach algebra with an involution, such that:

$$||a^\star a|| = ||a||^2.$$

Definition II.2. *Let A be an $\star$-algebra and (A, N) a fuzzy normed algebra. The fuzzy normed algebra (A, N) is called a fuzzy normed $\star$-algebra if*

$$N(a^\star, t) = N(a, t) \quad \forall a, a^\star \in A, \forall t \in \mathbb{R}^+.$$

A complete fuzzy normed $*$-algebra is called a fuzzy Banach $*$-algebra.

Definition II.3. *Let (A, N) be an fuzzy Banach $*$-algebra. The (A, N) is a called a fuzzy C^*-algebra if*

$$N(a^\star a, st) = N(a^\star, s)N(a, t) \quad \forall a, a^\star \in A, \forall s, t \in \mathbb{R}^+.$$

In this article we assume that m_0 is natural number. Also assume that

$$\mathbb{T}^1 = \{z \in \mathbb{C} : |z| = 1\}$$

and

$$\mathbb{T}^1_{\frac{1}{m_0}} := \{e^{i\theta}; 0 \leq \theta \leq \frac{2\pi}{m_0}\}.$$

Moreover, we suppose that (A, N) is fuzzy Banach algebra. For a given mapping $f : A \to A$, we define

$$D_{\lambda\gamma} f(a, b) = \bar{\lambda} f(\frac{a + \gamma b}{2}) + \bar{\lambda} f(\frac{a - \gamma b}{2}) - f(\lambda a) \tag{3}$$

$\forall a, b \in A, \forall \lambda, \gamma \in \mathbb{C}$.

Theorem II.1. *Let $p \in (0, \frac{1}{2})$ and $\theta \in [0, \infty)$ be real numbers. Suppose that $f : A \to A$ satisfies satisfying*

$$N(D_{\lambda\gamma} f(a, b), t) \geq \frac{t}{t + \theta(||a||^p ||b||^p)}, \tag{4}$$

$$N(f(ab) - f(b)f(a)) \geq \frac{t}{t + \theta(||a||^p ||b||^p)}, \tag{5}$$

$$N - \lim_{k \to \infty} 2^{-k} f(2^k (N - \lim_{k \to \infty} 2^{-k} f(2^k a))) = a, \tag{6}$$

for all $a, b \in A$, all $t > 0$ and all $\lambda, \gamma \in \mathbb{T}^1_{\frac{1}{m_0}}$.
Then there exists a unique involution $H : A \to A$ such that

$$H(a) := N - \lim_{k \to \infty} \frac{1}{2^k} f(2^k a)$$

and

$$N(f(a) - H(a), t) \geq \frac{(1 - 2^{2p-1})t}{(1 - 2^{2p-1})t + 2^{2p-1}\theta ||a||^{2p}} \tag{7}$$

$\forall a \in A, \forall t > 0$.
Also, if

$$N([N(f(a), t) - N(a, t)]a, t) \geq \frac{t}{t + \theta ||a||^{2p}} \tag{8}$$

$\forall a \in A, \forall t > 0$, *then (A, N) is a fuzzy banach $*$-algebra.*

Moreover, if

$$N([N(f(a)a,st)-N(f(a),s)N(a,t)]a,t) \geq \frac{t}{t+\theta||a||^{2p}}, \tag{9}$$

$\forall a \in A, \forall s,t>0$.
Then A is a C^*-algebra with involution f.

Proof. Substituting a and b with $2a$, λ by 1 and γ by 0 in (4), we get

$$N(\frac{1}{2}[f(2a)-2f(a)],\frac{1}{2}t) \geq \frac{t}{t+\theta||2a||^P||2a||^P},$$

$\forall a \in A, \forall t>0$.
Thus

$$N(\frac{1}{2}f(2a)-f(a),t) \geq \frac{t}{t+\theta 2^{2p-1}||a||^{2P}}, \tag{10}$$

$\forall a \in A, \forall t>0$.
Concider the set $S := \{g : A \to A\}$ and introduce the generalized metric on S:

$$d(g,h) = = \inf\{\delta \in \mathbb{R}^+ : N(g(a)-h(a),\delta t) \geq \frac{t}{t+\theta||a||^{2p}}, \quad \forall a \in A, \forall t>0\},$$

where, as usual, $\inf \theta = +\infty$.
It is easy to show that (S,d) is complete.
Now we define mappings $J : S \to S$ by

$$Jg(a) := \frac{1}{2}g(2a) \qquad \forall a \in A.$$

Let $g,h \in S$ be given such that $d(g,h) \neq +\infty$.
Then for some $\epsilon > 0$

$$N(g(a)-h(a),\epsilon t) \geq \frac{t}{t+\theta||a||^{2p}} \qquad \forall a \in A, \forall t>0.$$

Therefore

$$\begin{aligned} &N(Jg(a)-Jh(a),2^{2p-1}\epsilon t) \\ &= N(\frac{1}{2}g(2a)-\frac{1}{2}h(2a),2^{2p-1}\epsilon t) \\ &= N(g(2a)-h(2a),2^{2p}\epsilon t) \\ &\geq \frac{2^{2p}t}{2^{2p}t+\theta||2a||^{2p}} = \frac{t}{t+\theta||a||^{2p}}, \end{aligned}$$

for all a in A and all $t>0$.
If we define $L = 2^{2p-1} < 1$, then using the definition of d metric, we can conclude that $d(Jg,Jh) \leq L\epsilon$.
This means that

$$d(Jg,Jh) \leq Ld(g,h) \quad \forall g,h \in S.$$

It follows from (10) that $d(f,Jf) \leq L$.
Now, It follows from Theorem $I.1$ (i) that there exists a function $H : A \to A$ which is a fixed point of J i.e, $H(2a) = 2H(a) \quad \forall a \in A$, such that $\lim_{n\to\infty} d(J^n f,H) = 0$. Therefore, it can be concluded that

$$N-\lim_{n\to\infty}\frac{1}{2^n}f(2^n a) = H(a) \qquad \forall a \in A.$$

Since the integer n_0 of theorem $I.1(ii)$ is 0, then $H \in X^*$, which:

$$X^* = \{g \in S : d(f,g) < \infty\}.$$

By theorem $I.1(iii)$ and $d(f,Jf) \leq L$ we obtain

$$d(f,H) \leq \frac{1}{1-L}d(f,Jf) \leq \frac{L}{1-L},$$

i.e, the inequality (7) is true for all $a \in A$.
Suppose $\lambda = \gamma = 1$ and substituting a with $2^k a$ and b with $2^k b$ in (4)

$$N(\frac{1}{2^k}f(2^k\frac{a+b}{2})+\frac{1}{2^k}f(2^k\frac{a-b}{2})-\frac{1}{2^k}f(2^k a),\frac{1}{2^k}t) \geq \frac{t}{t+\theta(||2^k a||^p||2^k b||^p)},$$

thus

$$N(\frac{1}{2^k}f(2^k\frac{a+b}{2})+\frac{1}{2^k}f(2^k\frac{a-b}{2})-\frac{1}{2^k}f(2^k a),t) \geq \frac{2^k t}{2^k t+2^{2kp}\theta(||a||^p||b||^p)},$$

for all $a,b \in A$ and for all $t>0$.
Since

$$\lim_{k\to\infty}\frac{2^k t}{2^k t+2^{2kp}\theta(||a||^p||b||^p)} = 1,$$

then we have

$$N(H(\frac{a+b}{2})+H(\frac{a-b}{2})-H(a),t) = 1,$$

$\forall a,b \in A, \forall t>0$.
So the above equation shows H is a cauchy additive.
If $b=0$ in (4), then we have

$$N(\bar{\lambda}f(\frac{a}{2})+\bar{\lambda}f(\frac{a}{2})-f(\lambda a),t) \geq \frac{t}{t+0}$$

for all $a \in A$, $t>0$ and $\lambda \in \mathbb{T}^1_{\frac{1}{m_0}}$.
Then we have

$$N(2\bar{\lambda}H(\frac{a}{2})-H(\lambda a),t) = 1,$$

for all $a \in A$, $t>0$ and $\lambda \in \mathbb{T}^1_{\frac{1}{m_0}}$.
It follows by last equation and additivity of H that

$$H(\lambda a) = \bar{\lambda}H(a) \quad \forall \lambda \in \mathbb{T}^1_{\frac{1}{m_0}}.$$

Now, using the above materiales and the technique used in the [1], it can be easily shown that $H : A \to A$ is conjugate $\mathbb{C}$-linear.
By (5)

$$N(\frac{1}{4^k}f(4^k ab)-\frac{1}{2^k}f(2^K b).\frac{1}{2^k}f(2^k a),\frac{1}{4^k}t) \geq \frac{t}{t+\theta(||2^k a||^p||2^k b||^p)},$$

$\forall a,b \in A$ and $\forall t>0$.
Since

$$\lim_{k\to\infty}\frac{4^K t}{4^k t+4^{kp}\theta(||a||^p||b||^p)} = 1,$$

for all $a,b \in A$ and all $t>0$, therefore

$$N(H(ab)-H(b)H(a),t) = 1 \qquad \forall a,b \in A, \forall t>0.$$

Thus, we get for all $a, b \in A$ the equality

$$H(ab) = H(b)H(a).$$

On the other hand by (6)

$$H(H(a)) = N - \lim_{k\to\infty} 2^{-k} f(2^k(N - \lim_{k\to\infty} 2^{-k} f(2^k a))) = a,$$

for all a in A.
Hence $H : A \to A$ is an involution satisfying (7).
In addition, we must prove the uniqueness of H.
In fact, assume the existence of another such involutin H satisfies (7); hence, $H'(2^k a) = 2^k H'(a)$, according to (7),

$$N(\frac{1}{2^k} f(2^k a) - \frac{1}{2^k} H'(2^k a), t) = N(f(2^k a) - H'(2^k a), 2^k t) \geq \frac{(1-L)2^k t}{(1-L)2^k t + L\theta(||2^k a||^{2p})}.$$

Thus

$$N(\frac{1}{2^k} f(2^k a) - \frac{1}{2^k} H'(2^k a), t) \geq \frac{(1-L)2^k t}{(1-L)2^k t + L2^{2kp}\theta(||a||^P)}.$$

Because again

$$\lim_{k\to\infty} \frac{(1-L)2^k t}{(1-L)2^k t + L2^{2kp}\theta(||a||^P)} = 1,$$

for all $a \in A$, all $t > 0$, then

$$N(H(a) - H'(a), t) = 1 \qquad \forall a \in A, \forall t > 0.$$

Therefore, $H(a) = H'(a)$ for all $a \in A$.
Now, suppose H satisfies (8). Then we have

$$N([N(\frac{1}{2^k} f(2^K a), \frac{1}{2^k} t) - N(a, \frac{1}{2^k} t)]a, \frac{1}{2^k} t) \geq \frac{t}{t + \theta||2^k a||^{2p}}.$$

Thus

$$N([N(\frac{1}{2^k} f(2^k a), t) - N(a, t)]a, t) \geq \frac{2^k t}{2^k t + \theta||2^k a||^{2p}},$$

$\forall a \in A, \forall t > 0$.
Since

$$\lim_{k\to\infty} \frac{2^k t}{2^k t + \theta||2^k a||^{2p}} = 1,$$

it can be concluded that $N(H(a), t) = N(a, t))$, forall $a \in A$ and $t > 0$. Therefore, (A, N) is a fuzzy banach $*$-algebra.
Now, suppose H satisfies (9). Then we have

$$N([N(\frac{1}{2^k} f(2^k a)a, \frac{1}{2^{2k}} st) - N(\frac{1}{2^k} f(2^k a), \frac{1}{2^k} s)N(a, \frac{1}{2^k} t)]a, \frac{1}{2^k} t) \geq \frac{t}{t + \theta||2^k a||^{2p}}.$$

Thus

$$N([N(\frac{1}{2^k} f(2^k a)a, st) - N(\frac{1}{2^k} f(2^k a), s)N(a, t)]a, t) \geq \frac{2^k t}{2^k t + 2^{2kp}\theta||a||^{2p}} \quad \forall a \in A, \forall s, t > 0.$$

Because again

$$\lim_{k\to\infty} \frac{2^k t}{2^k t + 2^{2kp}\theta||a||^{2p}},$$

then

$$N(H(a)a, st) = N(H(a), s)N(a, t) \quad \forall a \in A, \forall t > 0.$$

Therefore A is a C^*-algebra with involution $a^* = H(a)$ for all $a \in A$. □

Theorem II.2. *Let $p \in (0, 1)$ and $\theta \in [0, \infty)$ be real numbers. Suppose that $f : A \to A$ with $f(1) = 1$, satisfies*

$$N(D_{\lambda\gamma} f(a, b), t) \geq \frac{t}{t + \theta(||a||^p + ||b||^p)},$$

$$N(f(ab) - f(b)f(a)) \geq \frac{t}{t + \theta(||a||^p + ||b||^p)},$$

$$N - \lim_{k\to\infty} 2^{-k} f(2^k(N - \lim_{k\to\infty} 2^{-k} f(2^k a))) = a,$$

for $\lambda, \gamma \in \mathbb{T}^1_{\frac{1}{m_0}}$ and $a, b \in A$.
Then there exists a unique involution $H : A \to A$ such that

$$H(a) := N - \lim_{k\to\infty} \frac{1}{2^k} f(2^k a)$$

and

$$N(f(a) - H(a), t) \geq \frac{(1 - 2^{p-1})t}{(1 - 2^{p-1})t + 2^{p-1}\theta||a||^p},$$

$\forall a \in A, \forall t > 0$.
Also, if

$$N([N(f(a), t) - N(a, t)]a, t) \geq \frac{t}{t + \theta||a||^p},$$

$\forall a \in A, \forall t > 0$, then (A, N) is a fuzzy banach $*$-algebra. Moreover, if

$$N([N(f(a)a, st) - N(f(a), s)N(a, t)]a, t) \geq \frac{t}{t + \theta||a||^p},$$

$\forall a \in A, \forall s, t > 0$.
Then (A, N) is a fuzzy C^*-algebra with involution $a^* = H(a)$, for all $a \in A$.

Proof. The proof of Theorem II.2 is completely similar to the that of TheoremII.1 and thus it is omitted. This com-pletes the proof of the theorem.

REFERENCES

[1] M. Eshaghi Gordji, *Nearly involutions on Banach algebra. a fixed point approach*, Fixed Point Theory 14(2013), 117-124.

[2] D. H. Hyers, *On the stability of the linear functional equation*, Proc. Nat. Acad. Sci. USA, 27 (1941), 222-224.

[3] A. Najati, fuzzy stability of a generalized quadratic functional equation, Common. Korean Math. Soc., 25(2010), 405-417.

[4] C. Park, M. Eshaghi Gorgji, R. Saadati, *Random homomorphisms and random derivations in random normed algebras via fixed point method*, J. Inequal. Appl, 2012 (2012), 13 pp. 1.4

[5] Th. M. Rassias, *On the stability of the linear mapping in Banach spaces*, Proc. Amer. Math. soc, 72(1978), 297-300.

[6] S. M. Ulam, *Problems in Modern Mathematics*, chapter VI, Science Ed. Wiley, New York, 1940.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Intelligent detection of bone fractures using data mining and image processing methods

Zahra Moradi Pour, Mohammad Reza Yamaghani

Department of Computer of Engineering, Azad University, Lahijan, Iran, za.moradipour@gmail.com, O_yamaghani@liau.ac.ir

***Abstract*—Today, X-ray machines produce high-quality images for radiologists and physicians to interpret images and detection the disease. Bone fractures are a relatively common event. Each type may require specific treatment, and it is important to determine the type of bone fracture in the fracture treatment phase. It may be a little difficult for a doctor to detect all types of fractures from X-rays and treatment methods. Therefore, an effective and accurate solution is needed to accelerate this procedure. This paper uses data mining and image processing methods which creates a reliable system that automatically and intelligently detects bone fractures with high accuracy and classifies bone fractures using classification models. The results show that not only the presented approach reduces time, effort and cost features, but also according to the experiments the system is able to detect more than 95% of bone fractures, which is an acceptable result.**

***Index Terms* —Intelligent detection; Bone fracture; Data mining methods; Image processing; x-ray images; Classification models**

I. INTRODUCTION

Bones are the hardest tissue in the body but may be cracked or broken due to high pressure or force. Bone fractures can have a variety of causes, but the most common causes include injuries (accidents), pathological causes (osteoporosis), stress, cancer, or overuse [1-3]. In addition, bone fractures are one of the leading causes of death in middle-aged and elderly people [4]. Therefore, each of them may need a specific treatment and determining the type of bone fracture in the fracture treatment stage is very important [1], and in general, X-ray images are used to decide and detection a fracture or not a bone fracture. In addition to X-rays, computed tomography (CT) and magnetic resonance imaging (MRI) are also used in the diagnosis of bone fractures. However, because of its low cost, user-friendliness, and easier accessibility, bone fracture detection is based on X-ray images [5-7]. After X-ray imaging, the location, type and surface of the fracture are identified and based on them, the types of fractures are classified, which is the first step in fracture treatment [8]. The main purpose of bone fracture classification models is to guide the treatment process and predict the side effects that may occur during treatment [9]. Therefore, before treatment, it is very important to determine the type of bone fracture in order to choose the appropriate treatment. Many classification methods deviate from the objectives of classification. In addition, it is difficult for a radiologist or a physician to remember all types of fractures and to identify them. Therefore, identifying different types of fractures with the help of computer system is an important necessity [5]. A computer system is a diagnosis made by a radiologist from the output of a computer analysis that uses medical images as a second perspective when making a diagnosis. This interpretation of the image helps the radiologist to improve the accuracy and coordination of the radiological diagnosis as well as to reduce the time of reading the image. In fact, computer-based automatic detection is based solely on computer algorithms [10]. Fortunately, developments in the computer's perspective have demonstrated the ability of computers to interpret medical images to the opinions of others. Because good algorithms search in a consistent manner, they may one day be able to distinguish between natural and unnatural lesions, despite their similar properties. This type of computer technology is a promising and efficient resource to help the radiologist improve diagnostic performance [11].

In the field of diagnosis, the accuracy and speed of diagnosis is of particular importance because it can sometimes be done incorrectly or with a long delay that causes irreparable damage. In this paper, using data mining techniques and methods and X-ray image processing, **a reliable system has been created that overcomes the existing limitations and automatically and, intelligently detects bone fractures Which can help the doctor or expert manpower to detection faster and more accurately and prevent irreparable damage and get the best result.** Since the existing systems do not have high

reliability and their comprehensiveness is low, in this research, an attempt is made to design and implement an automated system for intelligent bone fracture detection with acceptable performance.

The rest of the article is organized as follows. The definition of basic concepts is given in the second part. The third section provides information about designing the proposed method. The database is tested and the results of the performance evaluation of the proposed method compared to other methods are reported in Section IV. Finally, conclusions and references are included.

II. BACKGROUND

A. Bone and bone fracture

There are different types of bones and in addition there are different types of fractures that can vary from simple to complex types depending on the location and complexity. In this study, the focus is on the types of leg bone fractures (fibula and tibia) (Figure 1) which by analyzing medical images and then using the proposed methods, the type of fracture is diagnosed and the types of fractures are classified and finally leg bone fractures (fibula and tibia) are detected automatically and intelligently.

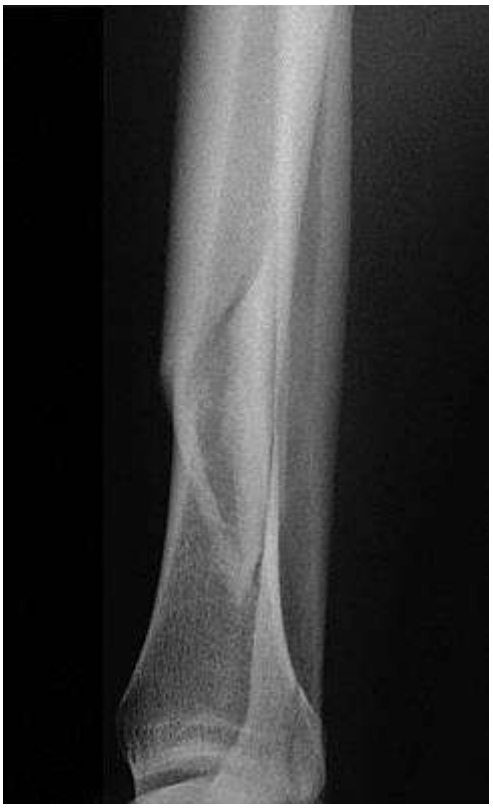


Figure 1. Leg bone fracture

B. Medical images

Today, medical images is a very important field in the detection disease. X-rays, computed tomography (CT), and magnetic resonance imaging (MRI) images are used to detection a variety of diseases. Many diagnoses rely heavily on an experienced human resource (specialist physician or radiologist). Bone fracture is a disease that is diagnosed by a human specialist (specialist doctor or radiologist) in which doctors use X-ray images because of the low cost, high speed, wider and easier access than CT and MRI images. They also use X-ray images to detect leg bone fractures (fibula and tibia) because of the lower X-ray dose than CT. The system proposed in this paper analyzes the medical image (X-ray) of the leg bone (fibula and tibia) which is processed using X-ray image and the following describes how to process the X-ray image.

III. MATERIAS AND METHODS

A. preprocessing

This step is done before the main steps in image processing. In this study, the contrast is improved and the image distortions are removed so that the objects in the image can be displayed more clearly and then the RGB images become gray images. The general objectives of this operation can be considered as image enhancement and removal of unnecessary data from the image. By performing proper preprocessing operations on images, detection accuracy and classification efficiency are improved [12].

B. Local binary pattern (LBP)

LBP is one of the features used for tissue classification. An LBP operator is a technique that converts an image into integer tags or an array of an image that provides a description of the microscopic appearance of the image. Statistics or image tags, usually histograms, are subsequently used to perform more detailed image analysis. In fact, it works very well for black and white images. The LBP feature offers promising results and performs better among the available methods. The LBP method is applied directly to the black-and-white image so that it can direct the image for proper classification. The LBP-based feature has poor performance in describing the color of an image, which in the case of a color image must first be turned gray before calculating the LBP. LBP captures the texture information of a color image to turn it into a gray space. In fact, LBP extracts image texture information by considering the pixel values around it and its neighbors. LBP captures the image information of an image to extract a corresponding image descriptor, and the LBP-based feature simply compares the center value and compares the pixel value (current processing pixel) with the neighboring pixel values in black and white. The results of the threshold bit values are stored in a matrix [13-15]. In fact, it can be said that LBP is considered as a threshold value in substituting the value of each pixel of the gray image based on the 8-bit binary code. The selected value of the central pixel, which is used as the threshold, is then selected to label the neighboring pixels by examining the variations between the gray scale values of the threshold. This method results in a new image based on a new LBP descriptor. This method is done by converting the binary code to a decimal number to the right of the threshold. Calculations start continuously clockwise. The function used as the threshold is (S(x)). Many values less than 0 are assigned 0, while values greater than 0 are assigned 1. Next, to produce an LBP image, all zeros and ones of the binary code are

converted to decimal values. The traditional LBP process is presented in Figure 2. This process is based on the number of conversions between zero and one in the patterns. The same aspects of LBP that are the main concept of LBP are obtained. The result of traditional LBP is classified into two main types of LBP values. The first type of LBP uniform consists of less than two or two bit conversions from one to zero or vice versa. For example, zero conversion (0,000,000) and two conversions (011111110). The second type of LBP is non-uniform LBP, which has more than two-bit connectors, such as four transitions (11101101) and six converters (01011011). LBP has a relationship between each uniform pattern and the result label, as shown in Figure 3.Therefore, pattern mapping can be done based on the number of separate result tags.

Different thresholds are used to identify uniform operators. The traditional threshold is the mean threshold and is used to identify the first LBP uniform operator. However, the uniform LBP properties identified by the mean threshold are affected by spot noise. This type of noise is included as a descriptor, because sometimes important image information is provided by spot noise. The properties of the LBP uniform are identified using the mean value for the eight pixels associated with the center. However, spot noise does not affect the features that can be identified at this stage. The threshold value is calculated based on the average of each image cell (3 x 3). The mean is then used as the threshold instead of many of the pixel centers used, as shown in Figure 4 [15]. In this way, the feature extraction process is performed based on different uniform LBP operators, followed by dimension reduction, and after dimension reduction, useful features are given to the classifier to determine the correct percentage.

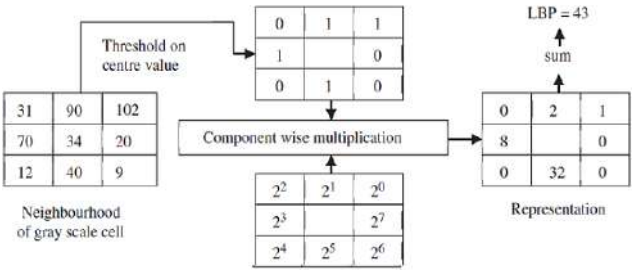


Figure 2. LBP method [15]

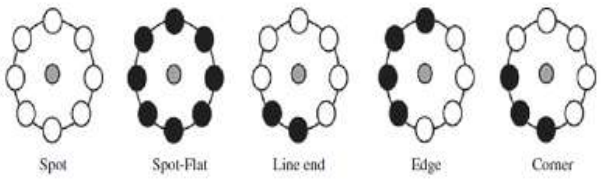


Figure 3. Initial identification of LBP [15]

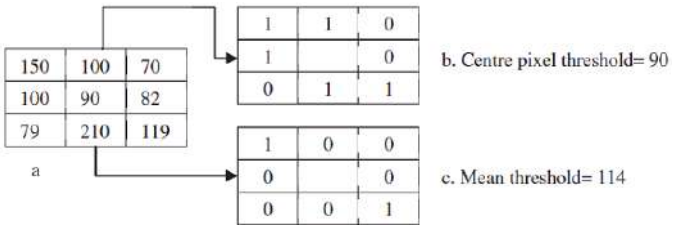


Figure 4. Average value for LBP [15]

C. *Principal component analysis (PCA)*

PCA is a well-known scheme for feature extraction and dimensionality, and is used to combine images [16]. Some of these features are misleading and some are extra and lead to double counting and therefore the dimension reduction is done to get useful features.

D. *Classification*

- K-nearest neighbors (KNN)

KNN is a simple and effective technique for classifying objects, which according to the closest examples of training in the feature space. Among training examples, class labels are used from the nearest neighbors to predict the test vector class. Therefore KNN meets the new points and according to the majority of votes obtained for K the closest score in the educational data, classifies them. In KNN, the Euclidean distance is often used as a distance metric to measure the similarity between two vectors (points) [17].

- Support vector machine (SVM)

SVMs are an effective method for solving classification and regression problems. SVM is particularly effective on data sets that are separated linearly, That is, where hiperplane H can be found to divide the samples into two classes so that the items in a class are (almost) completely on one side of H. Since there are an unlimited number of candidate hyperplanes that can be selected, SVM hyperplane chooses the H to maximize the distance to the nearest given score in both classes. This refers to maximizing the margin [18]. Good results are very efficient in feature extraction using support vector machines (SVM) and nearest neighborhood (KNN) with local binary patterns (LBP).

IV. DISUSSION AND RESULT

The main problem in accurately identifying bone on radiographs is changes in the shape and position of the patient, the exact location and division of the leg bone. Detection is made with a radiologist or orthopedist in the treatment of fractures. The correct diagnosis depends on the quality of the X-ray images and the physician's experience. Therefore, sufficient knowledge and experience in interpreting X-ray images is very important that manual extraction is time consuming and mentally prone to error. Traditional methods for detecting bone fractures from the image are difficult because of the unique differences in bone shape and the uncertainty of fracture shape.

Since a fixed shape of the bone specimens cannot be found. Therefore, we need a method that detects bone fractures according to the unique characteristics of the bone shape. In fact, a system that can automatically find the structure of the leg bone in an overview of X-ray images and detect its fracture. However, after examining the X-ray images and performing the preprocessing operation, we propose the automatic LBP method for extracting the characteristics of broken or healthy bone. LBP compares images pixel by pixel and acquires many features that not all of the features extracted by LBP are required and the PCA method is used to obtain a subset of features that using PCA, the dimension is reduced and appropriate features are selected from all the features to detection bone fractures. The extracted and selected properties from each image file are then saved, and this process is repeated for all fracture-related images and a training matrix and a test matrix were created for categorization and as well as making a matrix label for training so that the classifier can be taught with its help and the features were given to the KNN and SVM classifiers for training and testing and after training the category, the category was tested using the test matrix, and the test matrix labels were obtained, and by subtracting the output of the category and the test matrix labels, the number of correct diagnoses is displayed and then the desired detection percentage is obtained and displayed in the output.

Since the number of data was about 100 and was low and we used about 40 data from each class for training. Therefore, to test the system, we used the rest of the data of each class (about 10 images). The diagram of our proposed system is as follows:

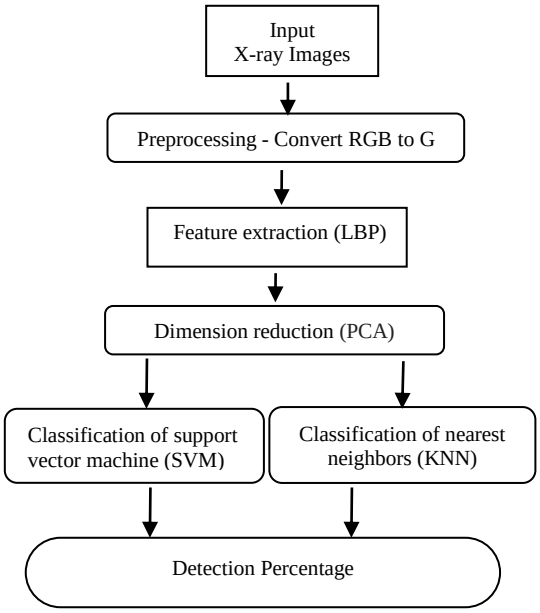


Figure 5. Flow chart of bone fracture detection

In this paper, the results consist of 100 samples of X-ray image data obtained from 100 radiology centers and all implementations in the proposed system were undertaken with the help of MATLAB software.

Experimental studies were also performed on a laptop with a Core i3 processor with 4 GB of memory and Windows 7 operating system. Also, for the first time, the LBP feature extraction method has been performed to identify bone fractures using X-ray images and has given very good results and PCA dimensional reduce method and KNN and SVM classification have been used for the first time on X-ray images of the leg bone that Compared to other methods, acceptable results have been presented and the results are shown in Table I:

TABLE I. TABLE DETECTION OF PERCENTAGE EACH CLASSIFICATION

Detection Percentage	Fracture Classification
97.83%	SVM
95.65%	KNN

Another advantage of the proposed method is the ability to perform fracture classification operations in a reasonable period of time. For a standard computer, this course is sufficient even in clinical terms. Accordingly, the proposed system can detect the type of fracture in less than 2.5 seconds. Detection time (in seconds) is the time required to preprocess, extract the feature, reduce the dimension, and classify a new image. In fact, each experiment was repeated 50 times and averaged from the time of identification, according to which the KNN classification method takes 2.47 seconds and the SVM classification method 2.18 seconds to detect the type of fracture that Based on the results obtained in Figure 6, the system spends most of its time extracting the feature. Figure 6 shows the fracture classification time, which has improved the time and accuracy of leg bone fracture detection compared to other studies.

Figure 6. FRACTURE CLASSIFICATION TIME

CONCLUSION

Bone fractures are a relatively common occurrence. Each type may require a specific treatment, and it is important to determine the type of bone fracture in the fracture treatment phase. X-ray images are examined manually but are time consuming and prone to error. Therefore, correct detection and timely treatment are very important for the patient and misdetection or long delay leads to irreparable damage. Due to the importance of bone fracture detection, in this article, using data mining and X-ray image processing methods and LBP

extraction method and PCA dimension reduction, a reliable system has been created which automatically and intelligently detects all types of bone fractures with high accuracy and using KNN and SVM classification models, bone fractures are classified with very high accuracy and it reduces time, effort, cost and errors and the results are acceptable and it can automatically and intelligently detect bone fractures with great accuracy and efficiency.

REFERENCES

[1] Bayram F, Çakıroğlu M, DIFFRACT: DIaphyseal Femur FRActure Classifier SysTem, Biocybernetics and Biomedical Engineering, 2016, 157 – 171.

[2] Graham TJ, Bamberger HB, Calandruccio JH, Wiedrich TA Focht LM, Bone plate extender and extension system for bone restoration and methods of use thereof, 2014.

[3] Umadevi N, Geethalakshmi DS. A brief study on human bone anatomy and bone fractures. IJCES Int J Comput Eng Sci Vol 2011; 1: 93–104.

[4] Liu Q, Cui X, Chou Y, Abbod M, Lin J, Ensemble artificial neural networks applied to predict the key riskfactors of hip bone fracture for elders, Biomedical Signal Processing and Control, 2015, 146 -152.

[5] Gamage P, Xie SQ, Delmas P, Xu WL, Diagnostic radiograph based 3D bone reconstruction framework: application to the femur, Comput Med Imaging Graph 2011, 427–37.

[6] Al-Ayyoub M, AL-ZGHOOL D, Determining the Type of Long Bone Fractures in X-Ray Images, WSEAS Transactions on Information Science and Applications,2013.

[7] Rankine J, (iv) Imaging of foot and ankle disorders, Orthopaedics and Trauma, 2009, 412– 419.

[8] Müller ME, Koch P, Nazarian S, Schatzker J. The comprehensive classification of fractures of long bones. Berlin: Springer-Verlag; 1990. p. 106–15.

[9] Marolt D, Knezevic M, Novakovic GV. Bone tissue engineering with human stem cells. Stem Cell Res Ther 2010; 1:10–1.

[10] K. Doi, Current status and future potential of computer-aided diagnosis in medical imaging, British Journal of Radiology, 2005, S3-S19.

[11] C. J. Vyborny, M. L. Giger, and R. M. Nishikawa, Computer-aided detection and diagnosis, Radiologic Clinics of North America, 2000, pp. 725–740.

[12] Shi, Z. and He, L. Application of neural networks in medical image processing. Proceedings of the Second International Symposium on Networking and Network Security. 2010.

[13] Liu P, Guo JM, Chamnongthai K, Prasetyo H, Fusion of Color Histogram and LBP-based Features for Texture Image Retrieval and Classification, Information Sciences,2017 ,95-111.

[14] Singh VP, Srivastava R, Automated and effective content-based mammogram retrieval using wavelet based CS-LBP feature and self-organizing map, Biocybernetics and Biomedical Engineering,2018,90-105.

[15] Zeebaree DQ, Abdulazeez AM, Zebari DA, Haron H, Abdull Hamed HN, Multi-Level Fusion in Ultrasound for Cancer Detection Based on Uniform LBP Features, Computers, Materials & Continua, 2020.

[16] Patil U, Mudengudi U, Image fusion using hierarchical PCA, International Conference on Image Information Processing(ICIIP),2011,1-6.

[17] Bijalwan V, Kumar V, Kumari P, Pascual J, KNN based Machine Learning Approach for Text and Document Mining, International Journal of Database Theory and Application,2014,61-70.

[18] Camlica Z, Tizhoosh HR, Khalvati F, Medical Image Classification via SVM using LBP Features from Saliency-Based Folded Data, International Conference on Machine Learning and Applications,2015,1-5.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Intelligent Transmitter : Analysis of Effective parameters on Sensor Response of Gas Transmitter to Enhancement Measurment Accuracy by Intelligent Corrective Model Based on Artificial Neural Network

Mohammad Mohseni Ahad[1], Touba Hamoule[2]

[1] *Department of Automation and Instrumentation Petroleum University of Technology Ahvaz, Iran.*
Mohsenieee2020@gmail.com
[2] *Department of Basic Sciences and Language Petroleum University of Technology Ahvaz, Iran.* t.hamoule@put.ac.ir

***Abstract*— As we may know, H2S is a toxic and dangerous gas that mainly can be found in oil fields, drilling rigs, gas separators, petrochemicals etc. It is so dangerous that if it exceeds a specified amount, it will cause physical and respiratory complications or death in some cases. For the time being, a transmitter is used to detect and measure the concentration of H2S gas which the most significant part of it, is the gas sensor. We used an electrochemical sensor to construct the transmitter. Neural networks have been used to investigate the effect of environmental parameters such as temperature and humidity. The network consists of an input layer, a hidden layer and an output layer. The results show that the output of the neural network is well able to follow the actual output. Therefore, changes in temperature and humidity affect the response of the hydrogen sulfide gas transmitter, and this change reduces the accuracy measured by the device. Furthermore, high-precision hydrogen sulfide detection sensors are generally expensive. This finding has important implications for developing robust gas sensors. By using the achieved relationship and considering the effect of temperature and humidity changes, the accuracy of the low-cos sensors can be greatly increased and the cost of producing a hydrogen sulfide transmitter can be greatly reduced.**

***Index Terms* — Artificial Neural Network, Intelligent Transmitter, Sensor response, Temperature and Humidity Effect**

I. INTRODUCTION

Hydrogen Sulfide is a toxic and dangerous gas which its mixture with air is flammable. It is so dangerous that if it exceeds a specified limit, it will cause respiratory disorders and death in some cases. Hydrogen Sulfide can be found in petrochemical industries, gas separators and deep or shallow wells. In oil fields, in different stage of well site, hydrogen sulfide can be found. In drilling stage, the source and manner of H2S is formation that have gases come with mud to above the ground and the leaking position is bell mouth and shale. In other stage, the source and manner of H2S is formation that have gases through finally cement casing and the leaking position is wellhead. In Testing stage, the source and manner of H2S is output fluid leaking and the leaking position is wellhead and extractor. In Production stage, the source and manner of H2S is chemical additive decaying and the leaking position is oil tank cover[1-3].

In order to obtain the safety of work place, transmitters are used to measure and control the concentration of hydrogen sulfide gas. A gas detector is a device used to measure gases in the environment and is known as part of the workplace safety system. This device also measures gas leakage and transfers it to the control room to perform the required operations in the face of gases. Their function is to generate an alarm when the gas concentration exceeds a certain level and buying them time to evacuate. The reason for the importance of these devices is that some gases are toxic and dangerous and if they exceed a certain level, they are harmful to human health. These devices are widely used in industry such as drilling rigs. Gas detectors can be classified in terms of the sensors used in them as well as in terms of installation capability. semiconductors, oxidation, catalytic, photoionization, infrared, portable devices and fixed gas detectors are types of gas detectors. [1, 2] Figure 1 illustrate the location of transmitter installation.

Figure 1Location of transmitter installation

So far, a lot of work has been done to manufacture gas transmitters, and many companies have built this device. However, among the device produced, there have been little discussion about the influence of temperature and humidity on transmitter output response and there are few cases that consider the effect of environmental factors on the output response measured by the transmitter and sensor and it is a major problem with this kind of devices. However, research and experiments have shown that ambient parameters such as temperature and humidity affect the output of the sensors (depending on the type of sensor structure) and the output of the transmitters. Recently, researchers have shown an increased interest in this issue. This concept has recently been challenged by Nenova studies demonstrating the effect of ambient factors on sensor output response. In 2013, Nenova, et.al. studied the efficacy of temperature and humidity on the amount of methane gas with metal oxide gas sensors and concluded that ambient factors such as temperature and humidity have substantial impact on metal oxide gas sensor characteristics.[3] However, far too little attention has been paid to analysis the efficacy of T and RH on transmitter output response and the most studies have only been carried out in a small number of areas such as data communication and multi gas sensor. In 2018, pang et.al. analysis the efficacy of T and RH on output response of electrochemical gas sensors detecting CO, NO, NO2, OX. They concluded that the T and RH have changing effect on the response of sensor in active and reference electrode. For NO gas, a linear relationship was found between T and RH and reference electrode. while other sensors have lower effect from ambient parameters.[4] This paper will focus on and construct a hydrogen sulfide gas transmitter and to investigate the effect of temperature and humidity parameters on the hydrogen sulfide sensor response with artificial neural network.

II. METHOD

A. Different Sensor Technologies

To date various methods have been developed and introduced to measure the amount of gas concentration. The more significant part of a transmitter is gas sensor which it plays a vital rule in transmitter designing. A variety of methods and sensors are used to detect gas concentration. Each has its advantage and drawbacks. Surface acoustic wave could be placed in harsh and rotating parts, detect nerve and blister agents and has a high response time. Catalytic gas sensor measures flammability of gases and requires air or oxygen to work. Optical gas sensors have a very wide monitoring area, easy operation in absence of oxygen and affected by ambient light interference. Thermal gas sensors have an easy operation in absence of oxygen and have reaction due to heating wire. Electrochemical gas Sensors have short response time, can measure toxic gases in relatively low concentrations and they can detect wide range of gases. Infrared gas sensors can be used in inert atmosphere but not all gases have IR absorption. Semiconductor gas sensors works well at constant high humidity condition and have an high response time.[1, 5-7]

Different studies have considered different sensors in a variety of ways. Considering the danger of hydrogen sulfide gas and the importance of its immediate detection, response time is the principle parameter in choosing a sensor. As shown in figure 4 and can be seen from table 1, electrochemical sensors has less response time compare to others.[8-10]

Table 1Comparison of all gas detection sensors

No	Sensor Type	Advantage	Disadvantage
1	Electrochemical	Short Response Time Measures toxic gases in relatively low concentrations. Wide range of gases can be detected	Advanced monitoring technique used
2	Infrared	Can be used in inert atmosphere	Not all gases have IR absorption.
3	Semiconductors	works well at constant high humidity condition.	Susceptible to contaminants and changes due to environmental conditions. High Response Time
4	Thermal	Robust but simple construction. Easy to operate in absence of oxygen.	Reaction due to heating wire.
5	Surface Acoustic Wave	Could be placed in harsh and rotating parts	High Response Time
6	Catalytic	measures flammability of gases low cost technology	Requires air or oxygen to work. Can be poisoned by lead, chlorine and silicones
7	Optical	Easy to operate in absence of oxygen. Not affected by electromagnetic interference	Affected by ambient light interference

Mainly, the use of gas detection sensors has brought so many opportunities to control infected environments. Nevertheless, there are serious issues about the quantity of the data acquired from sensors. Studies indicate that environmental parameters effect the output response gained from sensors and transmitters. Factors such as temperature and humidity effect the output response of sensors and transmitters and their measurement accuracy. This topic will be considering the efficacy of ambient factors such as T and RH on the output response of transmitter with hydrogen sulfide electrochemical sensor in variable circumstances and also on the measured concentration by the Sensor. [1, 11-14]

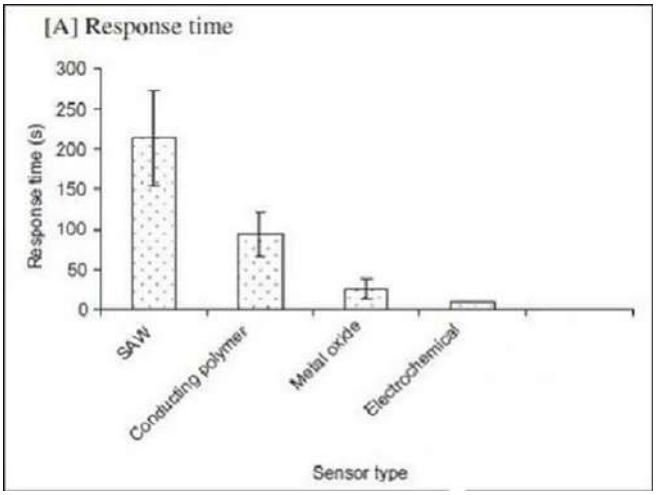


Figure 2 Comparison of response time of different electrochemical sensor

B. Electrochemical Sensor

The electrochemical sensor measures the gas concentration with oxidation reaction to generate positive or negative current through an external circuit. The main components of an electrochemical sensor include three electrodes which are: working, reference and counter. These components are combined with an electrolyte inside an electrochemical sensor. The upper part of the sensor has diffusion limiting hole through which the ambient air interacts with the sensor electrolyte. This results in electrochemical reaction. This reaction causes current to flow from "working" to "counter ". The value of this current is proportional to the amount of gas and is measured by an external electrochemical sensor circuit.

The basic of this sensors works is that the gas reacts with the sensor. A counter electrode is used to make a balance for working basic performance and also to create an equivalent current that fits to the intended gas. Such as those used in this device - ME3-H2S - some sensors also contain reference which is used for keeping sensitivity for the intended gas. Electrochemical sensors have admissible linearity. Previous studies confirm that in different conditions, the electrochemical sensor outputs will face changes and such changes is due to parameters like temperature and humidity. [4, 15, 16]

C. Design of Experiment Method (DOE)

Today, there is a lot of research being done on the design, test and manufacture of various equipment and materials. In order to increase the accuracy of the experiments and increase the efficiency of the proposed methods, this method should be introduced that are based on mathematical calculations and with which the experiment can be designed and modeled. Design of experiment creates an optimal experiment by considering the effect of different parameters and modeling the experiment and desired design. This method can be used to predict the process of testing and results obtained as well as to reduce the number of tests in physical and electrochemical experiments. In order to analysis the effect of T and RH parameters on response of hydrogen sulfide sensor, DOE-General Factorial approach it have been used. The figure illustrate that the data used is standard. [15] Here, data related to hydrogen sulfide gas of South Pars unit have been used. Due to the dangerous and toxic nature of this gas, its concentration and leakage are always controlled. In this study, the number of data is 80, of which 30 have been used for network training and the other 50 for network testing. Here, using the design of experiment (DOE) and general factorial methods, two modes are considered, which are in terms of the effect of temperature and humidity and not considering the effect of temperature and humidity on the sensor output and the amount of concentration measured. [17, 18]

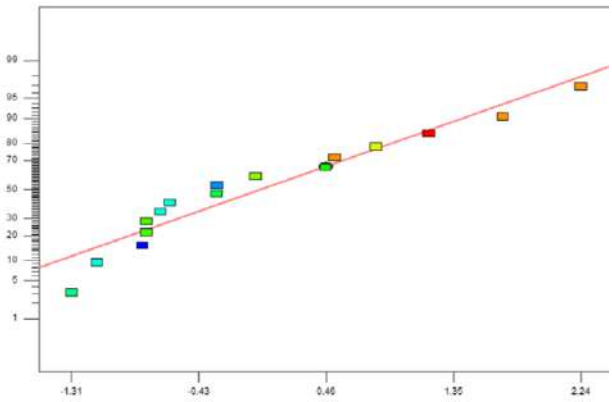


Figure 3Internally Studentized Residuals

D. Artificial Neural Network

Feed forward neural networks have been used to investigate the effect of environmental parameters such as temperature and humidity. This network consists of an input layer, a hidden layer and an output layer, the hidden layer of which consists of 6 neurons. The inputs of this network are the values of temperature, humidity and voltage of the sensor and its output is the value of the measured concentration. The total number of data is 80, of which 30 are used for training and 50 for testing. Data related to one of the wells in the oil-rich areas of South Pars.

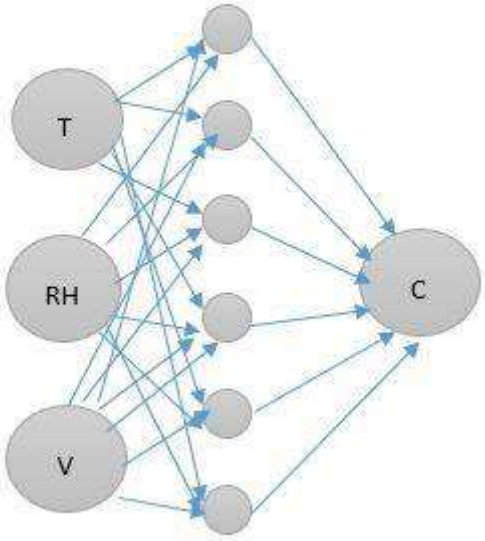


Figure 4 Structure of neural network

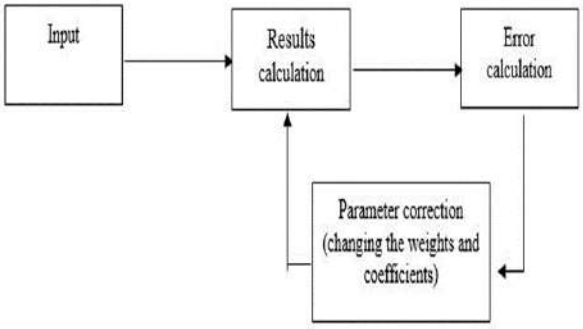


Figure 5Flowchart of training algorithm

III. Result and Discussion

By using this network, a model is obtained to correct measurement error due to environmental factors such as temperature and humidity. To acquire a model to forecast the process of the effect of the temperature and humidity on the output response of the transmitter, it was used quadratic polynomial curve fitting.

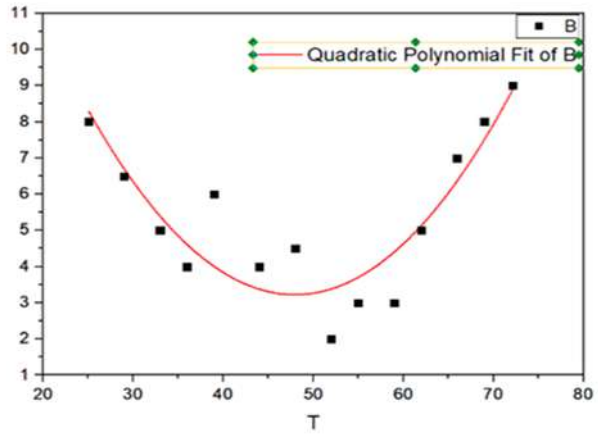


Figure 6H2S Sensor Response to the Change of Temperature

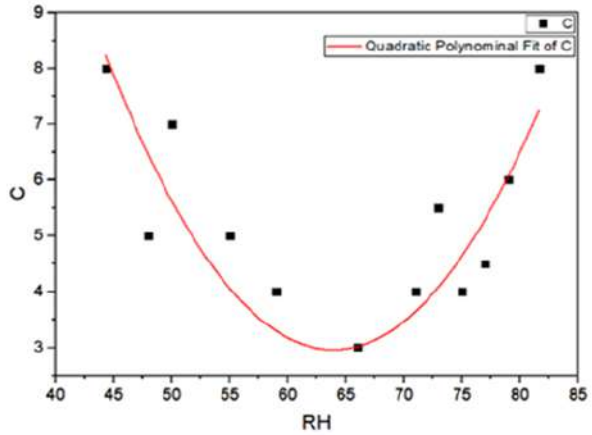


Figure 7 H2S Sensor Response to Change of Humidity Factor

$$Conc = (a*RH + b*T + k) * VAct$$

Table 2 sequential model sum of squares

Source	Sum of square	F-Value	P-Value
2FI vs Linear	0.18	0.043	0.8390
Quadratic vs 2FI	21.26	3.53	0.0690
Cubic vs Quadratic	9.32	0.67	0.6340

This formula of intelligent corrective model indicates that changes in each of the temperature and humidity factors have a parabolic relationship with the amount of response measured by the hydrogen sulfide transmitter.

Table 3 Analysis of variance

	Sum of squares	Mean of squares	P-Value	F-Value	Adeq Precision
Model	24.94	12.47	0.0325	4.51	5.668

The F-value of 4.51 point that the proposed model is significant. If the amount of P-value less than 0.05, the proposed model is significant. P-value indicate the error. As can been seen, the error is 3.25. Adeq precision measures the signal to noise ratio. If the obtained ratio is less than 4, the model is undesirable. As can been seen, the ratio is 5.668 and indicates an adequate signal.

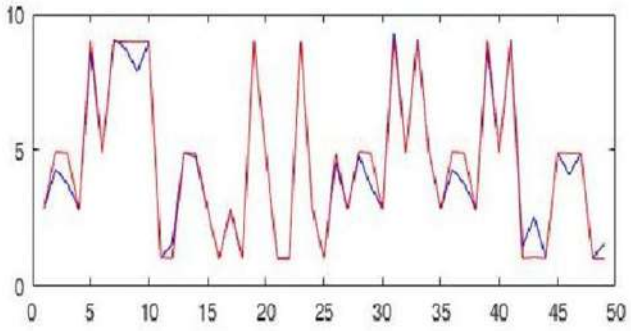


Figure 8 Result of intelligent corrective model

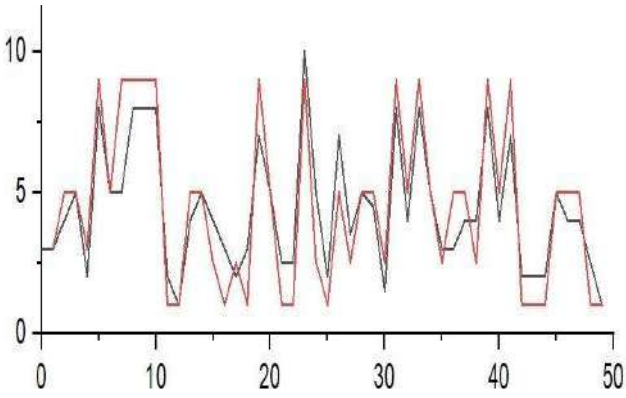


Figure 9Result without intelligent corrective model

By using this network, a model is obtained to correct measurement error due to environmental factors such as temperature and humidity. According to Figure 8, it can be seen that the obtained model has high accuracy than Figure 9. Its final prediction error is 0.16351 and the error in last iteration is 0.0066956 and three dimensional view of changes of T and RH are shown in figure 10. Axis X represents the number of test data and axis Y shows the amount of concentration. The red line represents the main data and the blue line also represents the model used.

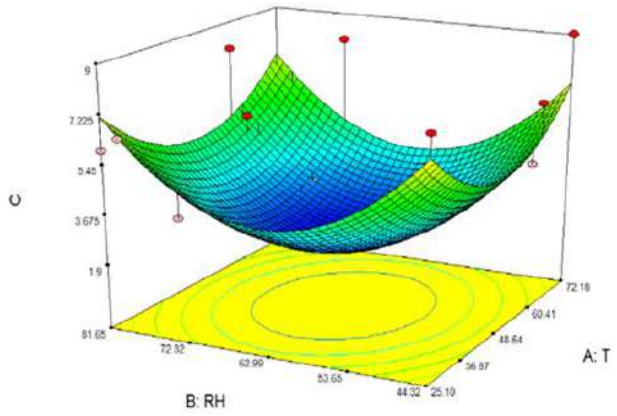


Figure 10 Three Dimension View of Effective Parameters

When considering the effect of temperature and humidity, it was observed that their changes affect the measured value. Therefore, using a neural network, intelligent corrective model was designed for it. In general, each sensor, according to the structural characteristics and the environmental condition in which it is located, has some changes in its output response. Using the proposed model and considering these changes, the output can be obtained with higher accuracy than the old fashioned sensors.

IV. Conclusion

In this study, the hydrogen sulfide transmitter performance under variable temperature and humidity conditions and its output response due to changes in the mentioned factors were investigated. This paper proposed an intelligent corrective model and the results of this model confirm many results of previous work in electrochemical gas sensors. This finding is in agreement with Wei (2018) findings which showed the effect of temperature and humidity on CO, NO and OX electrochemical gas sensor. According to the design of the experiment, it was observed that changes in temperature and humidity affect the working electrode of the sensor and as temperature and humidity change, the measured concentration also changes. Therefore, changes in temperature and humidity affect the response of the hydrogen sulfide gas transmitter, and this change reduces the accuracy measured by the device. Furthermore, high-precision hydrogen sulfide detection sensors are generally expensive. This finding has important implications for developing robust gas sensors. By using the achieved relationship and considering the effect of temperature and humidity changes, the accuracy of the low-cos sensors can be greatly increased and the cost of producing a hydrogen sulfide transmitter can be greatly reduced.

References

[1] S. K. Pandey, K.-H. Kim, and K.-T. Tang, "A review of sensor-based methods for monitoring hydrogen sulfide," *TrAC Trends in Analytical Chemistry*, vol. 32, pp. 87-99, 2012/02/01/ 2012, doi: https://doi.org/10.1016/j.trac.2011.08.008.

[2] B. Gao, X. Han, and H. Zhang, "Study on H2S Monitoring Technique for High Risk Wellsite," *Procedia Engineering*, vol. 45, pp. 898-903, 2012/01/01/ 2012, doi: https://doi.org/10.1016/j.proeng.2012.08.256.

[3] Z. Nenova and G. Dimchev, "Compensation of the impact of disturbing factors on gas sensor characteristics," *Acta Polytech. Hung*, vol. 10, no. 3, pp. 97-111, 2013.

[4] X. Pang, M. D. Shaw, S. Gillot, and A. C. Lewis, "The impacts of water vapour and co-pollutants on the performance of electrochemical gas sensors used for air quality monitoring," *Sensors and Actuators B: Chemical*, vol. 266, pp. 674-684, 2018/08/01/ 2018, doi: https://doi.org/10.1016/j.snb.2018.03.144.

[5] K. Vandana and C. Baweja, "Simmarpreet, and S. Chopra,"Influence of Temperature and Humidity on the Output Resistance Ratio of the MQ-135 Sensor,"" *Int. J. Adv. Res. Comput. Sci. Softw. Eng*, vol. 6, no. 4, pp. 423-429, 2016.

[6] S. Zampolli *et al.*, "An electronic nose based on solid state sensor arrays for low-cost indoor air quality monitoring applications," *Sensors and Actuators B: Chemical*, vol. 101, no. 1, pp. 39-46, 2004/06/15/ 2004, doi: https://doi.org/10.1016/j.snb.2004.02.024.

[7] R. Hosur, A. Rati, P. Dalawai, R. Gornal, and R. Patil, "A Survey on Automatic Detection of LPG Gas Leakage," in *2018 International Conference on Smart Systems and Inventive Technology (ICSSIT)*, 2018: IEEE, pp. 266-269.

[8] N. Afshar-Mohajer *et al.*, "Evaluation of low-cost electro-chemical sensors for environmental monitoring of ozone, nitrogen dioxide, and carbon monoxide," *Journal of occupational and environmental hygiene*, vol. 15, no. 2, pp. 87-98, 2018.

[9] Y. Liu, X. Ni, Y. Wu, and W. Zhang, "Study on effect of temperature and humidity on the CO2 concentration measurement," in *IOP Conference Series: Earth and Environmental Science*, 2017, vol. 81, no. 1: IOP Publishing, p. 012083.

[10] A. ŚPIEWAK and W. SAŁABUN, "A mobile gas detector with an Arduino microcontroller," *Int J Computer Technol Appl*, vol. 6, no. 4, pp. 636-641, 2016.

[11] K. Arunkumar, S. Surendar, S. Piraisudan, and G. Siva, "PREVENTION OF SEWAGE WORKERS FROM POISONOUS GAS," 2018.

[12] W. Mickelson, A. Sussman, and A. Zettl, "Low-power, fast, selective nanoparticle-based

hydrogen sulfide gas sensor," *Applied Physics Letters*, vol. 100, no. 17, p. 173110, 2012.
[13] H. Mahdavi, S. Rahbarpour, S.-M. Hosseini-Golgoo, and H. Jamaati, "Reducing the destructive effect of ambient humidity variations on gas detection capability of a temperature modulated gas sensor by calcium chloride," *Sensors and Actuators B: Chemical*, vol. 331, p. 129091, 2021.
[14] J. Dai *et al.*, "Ultrafast response polyelectrolyte humidity sensor for respiration monitoring," *ACS applied materials & interfaces*, vol. 11, no. 6, pp. 6483-6490, 2019.
[15] F. Sauli, "Gas detectors: achievements and trends," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 461, no. 1, pp. 47-54, 2001/04/01/ 2001, doi: https://doi.org/10.1016/S0168-9002(00)01165-7.
[16] S. Ghosh, M. S. Ilango, C. S. Prajapati, and N. Bhat, "Reduction of Humidity Effect in WO3 Thin Film - Based NO2 Sensor Using Physiochemical Optimization," *Crystal Research and Technology*, vol. 56, no. 1, p. 2000155, 2021.
[17] J. Miao, C. Chen, and J. Y. Lin, "Humidity independent hydrogen sulfide sensing response achieved with monolayer film of CuO nanosheets," *Sensors and Actuators B: Chemical*, vol. 309, p. 127785, 2020.
[18] M. Shooshtari, A. Salehi, and S. Vollebregt, "Effect of temperature and humidity on the sensing performance of TiO2 nanowire-based ethanol vapor sensors," *Nanotechnology*, vol. 32, no. 32, p. 325501, 2021.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

New results on MV-semimodules

1st Simin Saidi Goraghani and 2nd Rajab Ali Borzooei

1st *Department of Mathematics*, *Farhangian University*, Tehran, Iran, siminsaidi@yahoo.com
2nd *Department of Mathematics*, *Shahid Beheshti University*, Tehran, Iran, borzooei@sbu.ac.ir

***Abstract*— Recently, we introduced a new definition of MV-semimodules [8]. In this paper, we study A-ideals in MV-semimodules and Q-ideals in semirings. Then we verify the relationship between ideals of semirings and ideals of MV-algebras. Finally, we obtain some results on A-ideals that could be proved by the old definitions only under certain conditions.**

***Index Terms*— MV-*semi*ring, MV-algebra, MV-semimodule, Q-ideal, A-ideal.**

I. Introduction

The concept of semirings was introduced by H. S. Vandiver (1935). MV-semiring is an special ring that has recently been introduced. Nowadays, the theory of idempotent semirings has many applications in other fields. The concept of MV-algebras was defined by C.C. Chang as algebras corresponding to the Łukasiewicz Logic. In 2003, A. Di Nola presented the notion of MV-modules as a PMV-algebra multiplication in an MV-algebra. Then some researches worked on MV-modules (see, for instance, [5–7]), and some papers were written by that definition. In [2], Di Nola and C. Russo introduced the notion of MV-semimodules as an MV-semiring multiplication in an abelian monoid. In [8], we presented the notion of *new definition of MV-semimodules*. We considered multiplication of an MV-semiring in an MV-algebra in order to verify MV-modules and their ideals better and simpler than before.
In this paper, we study prime A-ideals in MV-semimodules, and we obtain some results that could be proved by the old definitions only under certain conditions. For example, we present some equivalent conditions for having prime A-ideals in MV-semimodules.

II. Preliminaries

In this section, we review the definitions that we need in the following sections.

Definition 1. *[2] An algebraic structure $(\Upsilon, \dot{+}, ., 0, 1)$ of type $(2, 2, 0, 0)$ is called a semiring if;*
(i) $(\Upsilon, \dot{+}, 0)$ is a commutative monoid,
(ii) $(\Upsilon, ., 1)$ is a monoid,
(iii) "$.$" distributes over "$\dot{+}$" from either side. Υ is said commutative if $t.h = h.t$, for every $t, h \in \Upsilon$. If Υ satisfies the equation $t \dot{+} t = t$, for every $t \in \Upsilon$, then Υ is called idempotent.

Consider $\emptyset \neq \Omega \subseteq \Upsilon$. Ω is called a left ideal of semiring Υ if it satisfies the following condition:
(1) if $t, h \in \Omega$, then $t \dot{+} h \in \Omega$,
(2) if $t \in \Omega$ and $\mu \in \Upsilon$, then $\mu.t \in \Omega$.

Semiring $(\Upsilon, \dot{+}, ., 0, 1)$ is called An MV-semiring, if it is commutative and additive idempotent. Moreover, there exists a map $\diamond : \Upsilon \to \Upsilon$ that satisfying the following conditions:
(i) $\mu.\nu = 0$ if and only if $\nu \leq \mu^{\diamond}$ (" $\leq$ " is naturally defined by means of $\dot{+}$),
(ii) $\mu + \nu = (\mu^{\diamond}.(\mu^{\diamond}.\nu)^{\diamond})^{\diamond}$,
for every $\mu, \nu \in \Upsilon$.

Definition 2. *[1] Algebra $V = (V, \oplus, \diamond, 0)$ of type $(2, 1, 0)$ is called an MV-algebra, if it satisfying the following equations*
$(MV1)$ $(V, \oplus, 0)$ is an Abelian monoid,
$(MV2)$ $(\mu^{\diamond})^{\diamond} = \mu$,
$(MV3)$ $0^{\diamond} \oplus \mu = 0^{\diamond}$,
$(MV4)$ $(\mu^{\diamond} \oplus h)^{\diamond} \oplus h = (h^{\diamond} \oplus \mu)^{\diamond} \oplus \mu$, for every $\mu, h \in V$.
If we define the constant $1 = 0^{\diamond}$ and operations $\odot$ and $\ominus$ by $\mu \odot h = (\mu^{\diamond} \oplus h^{\diamond})^{\diamond}$ and $\mu \ominus h = \mu \odot h^{\diamond}$, then
$(MV5)$ $(\mu \oplus h) = (\mu^{\diamond} \odot h^{\diamond})^{\diamond}$,
$(MV6)$ $\mu \oplus 1 = 1$,
$(MV7)$ $(\mu \ominus h) \oplus h = (h \ominus \mu) \oplus \mu$,
$(MV8)$ $\mu \oplus \mu^{\diamond} = 1$,
for every $\mu, h \in V$. We know that $(V, \odot, 1)$ is an Abelian monoid. If we define auxiliary operations $\vee$ and $\wedge$ on V by $\mu \vee h = (\mu \odot h^{\diamond}) \oplus h$ and $\mu \wedge h = \mu \odot (\mu^{\diamond} \oplus h)$, for every $\mu, h \in V$, then $(V, \vee, \wedge, 0)$ is a bounded distributive lattice. An ideal of MV-algebra V is a subset I of V, satisfying the following conditions: $(I1)$: $0 \in I$, $(I2)$: $\mu \leq h$ and $h \in I$ imply $\mu \in I$, $(I3)$: $\mu \oplus h \in I$, for every $\mu, h \in I$. A proper ideal I of V is a prime ideal of V if and only if $\mu \ominus h \in I$ or $h \ominus \mu \in I$ (or $\mu \wedge h \in I$ implies that $\mu \in I$ or $h \in I$), for every $\mu, h \in V$.

Definition 3. *[8] Let $A = (A, \dot{+}, ., 0, 1)$ be an MV-semiring, $V = (V, \oplus, \diamond, 0)$ be an MV-algebra, and $\phi : A \times V \longrightarrow V$ be defined by $\phi(t, v) = tv$, for every $t \in A$ and $v \in V$. Then V is called a (left) MV-semimodule over A or briefly an A-semimodule if for every $t, h \in A$ and $v_1, v_2 \in V$:*

$(SMV1)$ *if* $v_1 + v_2$ *is defined in* V*, then* $tv_1 + tv_2$ *is defined in* V *and* $t(v_1 + v_2) = tv_1 + tv_2$*;*
$(SMV2)$ $(t \dot{+} h)v_1 = tv_1 \oplus hv_1$*;*
$(SMV3)$ $(t.h)v_1 = t(hv_1)$*.*
V is a unitary A-semimodule if A has a unity 1_A for the product, that is
$(SMV4)$ $1_A v_1 = v_1$*, for every* $v_1 \in V$*.*

Proposition 1. *[2]* (i) *If* $\Upsilon = (\Upsilon, \oplus, \diamond, 0)$ *is an MV-algebra, then* $(\Upsilon, \vee, \circledcirc, 0, 1)$ *is a semiring.*
(ii) *If* $\Upsilon = (\Upsilon, \dot{+}, ., 0, 1)$ *is an MV-semiring, then* $(\Upsilon, \oplus, \diamond, 0)$ *is an MV-algebra, where* $\gamma \oplus \delta = (\gamma^\diamond.\delta^\diamond)^\diamond$*, for all* $\gamma, \delta \in \Upsilon$*.*

III. New results on MV-semimodules

Now, we study prime A-ideals in MV-semimodules and verify the conditions for having the prime A-ideals of MV-semimodules.

Definition 4. *Consider* $A = (A, \dot{+}, ., 0, 1)$ *is an MV-semiring,* $V = (V, \oplus, \diamond, 0)$ *is an MV-algebra, V is an A-semimodule, and Ψ is an ideal of V. Then Ψ is called an A-ideal of V if $(I4)$: $t\gamma \in \Psi$, for every $t \in A$ and $\gamma \in \Psi$. Moreover, Ψ is called a prime A-ideal of V, if $t\gamma \in \Psi$ implies that $\gamma \in \Psi$ or $t \in (\Psi : V) = \{t \in A : tV \subseteq \Psi\}$, for any $t \in A$ and $\gamma \in V$. The proper ideal ρ of A is called a prime ideal of A if $t.h \in \rho$ implies that $t \in \rho$ or $h \in \rho$, for any $t, h \in A$.*

Example 1. *Let* $A = \{e, s, t, u\}$ *and the operations "$\dot{+}$" and "." on A be defined as follows:*

$\dot{+}$	e	s	t	u
e	e	s	t	u
s	s	s	u	u
t	t	u	t	u
u	u	u	u	u

$.$	e	s	t	u
e	e	e	e	e
s	e	s	e	s
t	e	e	t	t
u	e	s	t	u

Consider the map $\diamond : A \longrightarrow A$ *such that* $e^\diamond = u$*,* $s^\diamond = t$*,* $t^\diamond = s$ *and* $u^\diamond = e$*. Then we can easily to show that* $(A, \dot{+}, ., e, u)$ *is an MV-semiring and* $(A, \oplus, \diamond, e)$ *is an MV-algebra, where* $\oplus = \dot{+}$*. Now, let the operation* $\phi : A \times A \longrightarrow A$ *be defined by* $\phi(t, h) = t.h = th$*, for every* $t, h \in A$*. We can easily to show that A is an A-semimodule. Also,* $\Psi = \{e, s\}$ *and* $\Gamma = \{e, t\}$ *are prime A-ideals of A and* $\Xi = \{e\}$ *is not a prime A-ideal of A. Moreover,* $I = \{e, s\}$ *is a prime ideal of A and* $J = \{e\}$ *is not a prime ideal of A.*

Proposition 2. *Let* $A = (A, \dot{+}, ., 0, 1)$ *be an MV-semiring, and Ω be an ideal of A as an MV-algebra. Then Ω is an ideal of A.*

Proof. By Proposition 1 (ii), we know that $(A, \oplus, \diamond, 0)$ is an MV-algebra, where $i \oplus j = (i^\diamond.j^\diamond)^\diamond$, for every $i, j \in A$. We have

$$i \dot{+} j = (i^\diamond.(i^\diamond.j)^\diamond)^\diamond = i \oplus (i \oplus j^\diamond)^\diamond.$$

Since $(i \oplus j^\diamond)^\diamond \leq j$, we have $(i \oplus j^\diamond)^\diamond \oplus i \leq j \oplus i$ and so $i \dot{+} j \leq i \oplus j \in \Omega$. It results that $i \dot{+} j \in \Omega$. Now, let $i \in \Omega$ and $r \in A$. Since $i.r \leq i \in \Omega$, we have $i.r \in \Omega$. Therefore, Ω is an ideal of A. □

Theorem 1. *If* $A = (A, \dot{+}, ., 0, 1)$ *is an MV-semiring such that* $i.i = i$*, for every* $i \in A$*, then for every ideal Ψ of A, we have:*
(i) *Ψ is an ideal of A as an MV-algebra;*
(ii) *If* $i.l^\diamond, l \in \Psi$*, then* $i \in \Psi$*, for any* $i, l \in A$*;*
(iii) *If Ψ is a prime ideal of A as an MV-semiring, then Ψ is a prime ideal of A as an MV-algebra;*
(iv) *Ψ is a prime ideal of A if and only if* $i.l^\diamond \in \Psi$ *or* $l.i^\diamond \in \Psi$*, for every* $i, l \in A$*.*

Proof. (i) By Proposition 1 (ii), we know that $(A, \oplus, \diamond, 0)$ is an MV-algebra, where $t \oplus h = (t^\diamond.h^\diamond)^\diamond$, for every $t, h \in A$. Since $i.i = i$, for every $i \in A$, We can easily to see that $t \oplus h \in \Psi$, for every $t, h \in \Psi$. Let $t \leq h$ and $h \in \Psi$, for some $t, h \in A$. We must show that $t \in \Psi$. Since $t \leq h$, we have $h^\diamond.t = 0 \in \Psi$. On the other hand, by $(SMV1)$ and $(SMV4)$, we have

$$t = t1 = t(h \oplus h^\diamond) = th \oplus th^\diamond = t.h \oplus t.h^\diamond \in \Psi.$$

Hence Ψ is an ideal of A as an MV-algebra.
(ii) Since $i.l^\diamond, l \in \Psi$, by (i), we have $i.l^\diamond \oplus l \in \Psi$. Also, since

$$i.(i.l^\diamond \oplus l)^\diamond = i.((i.l^\diamond)^\diamond.l^\diamond) = (i.l^\diamond).(i.l^\diamond)^\diamond = 0,$$

we have $i \leq i.l^\diamond \oplus l \in \Psi$ and so $i \in \Psi$.
(iii) Consider Ψ as a prime ideal of A (A is considered as an MV-semiring), and $i \wedge l \in \Psi$, for some $k.l \in A$. We must show that $k \in \Psi$ or $l \in \Psi$. Since

$$k.(k.l^\diamond)^\diamond = i.(i^\diamond \oplus l) = (i^\diamond \oplus (i^\diamond \oplus l)^\diamond)^\diamond = i \wedge l \in \Psi,$$

we have $i.(i.l^\diamond)^\diamond \in \Psi$ and so $i \in \Psi$ or $(i.l^\diamond)^\diamond \in \Psi$. Let $i \notin \Psi$. Then $(i.l^\diamond)^\diamond \in \Psi$. Now, we have $l \leq i^\diamond \oplus l = (i.l^\diamond)^\diamond \in \Psi$ and so by (i), we have $l \in \Psi$.
(iv) $(\Rightarrow)$ Consider Ψ as a prime ideal of A. Since $(i.l^\diamond).(l.i^\diamond) = 0 \in \Psi$, we have $i.l^\diamond \in \Psi$ or $l.i^\diamond \in \Psi$, for every $i, l \in A$.
$(\Leftarrow)$ Let $i.l^\diamond \in \Psi$ or $l.i^\diamond \in \Psi$, for every $i, l \in A$. Now, let $t.h \in \Psi$, $t \notin \Psi$ and $h \notin \Psi$. We consider $t.h^\diamond \in \Psi$. Then $t.(h \dot{+} h^\diamond) = t.h \dot{+} t.h^\diamond \in \Psi$ and so $h \dot{+} h^\diamond \in \Psi$. Now, we have

$$t.(h \dot{+} h^\diamond)^\diamond = t.(h^\diamond.(h^\diamond.h^\diamond)^\diamond) = (t.h^\diamond).(h^\diamond.h^\diamond)^\diamond \in \Psi$$

and so by (ii), we have $t \in \Psi$ that is a contradiction. Hence $t \in \Psi$ or $h \in \Psi$. □

In the follows, we construct the quotient MV-semiring by considering Q-ideals in MV-semirings.

Definition 5. *[3, 4] Let Υ be a semiring and I be an ideal of Υ. Then I is called a Q-ideal of Υ if there exists a subset*

Q of Υ such that:
(1) $\Upsilon = \bigcup\{d \dot{+} I : d \in Q\}$;
(2) *if* $d_1, d_2 \in Q$, *then* $(d_1 \dot{+} I) \cap (d_2 \dot{+} I) \neq \emptyset$ *if and only if* $d_1 = d_2$.

Now, let $\frac{\Upsilon}{I} = \{d \dot{+} I : d \in Q\}$. *Consider operations* $\boxplus$ *and* $\boxtimes$ *be defined on* $\frac{\Upsilon}{I}$ *as follows:*

$$(d_1 \dot{+} I) \boxplus (d_2 \dot{+} I) = d_3 \dot{+} I$$

where $d_3 \in Q$ *is a unique element such that* $d_1 \dot{+} d_2 \dot{+} I \subseteq d_3 \dot{+} I$.

$$(d_1 \dot{+} I) \boxtimes (d_2 \dot{+} I) = d_4 \dot{+} I$$

where $d_4 \in Q$ *is a unique element such that* $d_1 d_2 \dot{+} I \subset d_4 \dot{+} I$. *Then* $(\frac{\Upsilon}{I}, \boxplus, \boxtimes, 0 \dot{+} I)$ *is an* MV*-semiring related to* I.

Note. From now on, in this paper, we let A be an MV-semiring and V be an MV-algebra.

Definition 6. *Consider* V *as an* A*-semimodule and* Ψ *as an* A*-ideal of* V. *Then* Ψ *is said a torsion free* A*-ideal, if* $t\gamma = 0$ *implies that* $t = 0$ *or* $\gamma = 0$, *where* $t \in A$ *and* $\gamma \in \Psi$. *In case that* $V = \Psi$, V *is called a torsion free* MV*-semimodule on* A *(or torsion free* A*-semimodule).*

Example 2. (i) *In Example 1,* $I = \{e, s\}$ *and* $J = \{e, t\}$ *are torsion free* A*-ideals of* A.
(ii) *Let* $A = L_2 = \{0, 1\}$, $z \dot{+} f = min\{1, z + f\}$, *and the map* $\diamond : L_2 \longrightarrow L_2$ *be defined by* $z^\diamond = 1 - z$, *for every* $z, f \in L_2$, *where* $+, -, .$ *are ordinary operations in* $\mathbb{R}$. *Then it is routine to show that* $(L_2, \dot{+}, ., 0, 1)$ *is an* MV*-semiring. Also, let* $V = L_4 = \{0, \frac{1}{3}, \frac{2}{3}, 1\}$ *and operations* " $\dot{+}$ " *and* " $\diamond$ " *be defined on* L_4 *similar to* L_2. *Then it is routine to show that* $(V, \dot{+}, \diamond, 0)$ *is an* MV*-algebra. Now, let operation* $\phi : A \times V \longrightarrow V$ *be defined by* $\phi(z, f) = z.f = zf$, *for every* $z \in A$ *and* $f \in V$. *Hence we can see that* V *is an* MV*-semimodule. Also,* L_4 *is a torsion free* MV*-semimodule over* L_2.
(iii) *In Example 1,* A *is not a torsion free* MV*-semimodule.*

Example 3. *Let* V *be a unitary* A*-semimodule, and* Ψ *be an* A*-ideal of* V.
(i) *If* Ψ *is a torsion free* A*-ideal of* V, *then* $Ann_A(\Psi)$ *is a* Q*-ideal of* A. *By considering* $Q = (A - Ann_A(\Psi)) \cup \{0\}$, *we can show that* $Ann_A(\Psi)$ *is a* Q*-ideal of* A.
(ii) *If* Ψ *is a prime* A*-ideal of* V *such that* $a^\diamond m \in \Psi$ *implies* $am \in \Psi$, *for every* $a \in A$ *and* $m \in V$, *then* $(\Psi : V)$ *is a* Q*-ideal of* A. *It is routine to see that* $(\Psi : V)$ *is an ideal of* A. *Let* $Q = (A - (\Psi : V)) \cup \{0\}$. *We show that* $(\Psi : V)$ *is a* Q*-ideal of* A. *It is easy to see that* $A = \bigcup\{q \dot{+} (\Psi : V) : q \in Q\}$. *Now, we show that* $(q_1 \dot{+} (\Psi : V)) \cap (q_2 \dot{+} (\Psi : V)) = \emptyset$, *for every* $q_1, q_2 \in Q$, *where* $q_1 \neq q_2$. *Let*

$$(q_1 \dot{+} (\Psi : V)) \cap (q_2 \dot{+} (\Psi : V)) \neq \emptyset,$$

for some $q_1, q_2 \in Q$, *where* $q_1 \neq q_2$ *and there exists* $r \in (q_1 \dot{+} (\Psi : V)) \cap (q_2 \dot{+} (\Psi : V))$. *There are* 2 *cases*:
(I) *Let* $q_1 = 0$, $q_2 \neq 0$. *Then there exists* $t \in (\Psi : V)$ *such that* $r = q_2 \dot{+} t$, $rV \subseteq \Psi$, $tV \subseteq \Psi$ *and* $q_2 V \nsubseteq \Psi$. *So* $rm, tm \in \Psi$, *for every* $m \in V$ *and there is* $m_1 \in V$ *such that* $q_2 m_1 \notin \Psi$. *Now, since* $q_2 m_1 \leq q_2 m_1 \oplus t m_1 = r m_1 \in \Psi$, *we have* $q_2 m_1 \in \Psi$ *that is a contradiction. Hence* $q_2 = 0$ *and so* $q_1 = q_2$.
(II) *Let* $q_1, q_2 \neq 0$, *where* $r = q_1 \dot{+} t_1 = q_2 \dot{+} t_2$, *for* $t_1, t_2 \in (\Psi : V)$. *Then* $q_1 V \nsubseteq \Psi$, $q_2 V \nsubseteq \Psi$ *and there exists* $m \in V$ *such that* $q_1 m \notin \Psi$. *We consider two cases:*
(1) *If* $q_2 m \in \Psi$, *then*

$$\begin{aligned} q_1 m \leq q_1 m \oplus t_1 m &= (q_1 \dot{+} t_1)m = rm = (q_2 \dot{+} t_2)m \\ &= q_2 m \oplus t_2 m \in \Psi \end{aligned}$$

and so $q_1 m \in \Psi$ *that is a contradiction. So* $q_1 = q_2$.
(2) *Let* $q_2 m \notin \Psi$, *too. Then since* $q_1 \dot{+} t_1 \leq q_2 \dot{+} t_2$, *we have* $(q_2 \dot{+} t_2)^\diamond . (q_1 \dot{+} t_1) = 0$ *and so*

$$q_2^\diamond . (q_2 \dot{+} t_2)^\diamond . (q_1 \dot{+} t_1) = 0.$$

It results that $q_2^\diamond . (q_1 \dot{+} t_1) \leq q_2^\diamond . t_2$ *and so*

$$\begin{aligned} (q_2^\diamond . (q_1 \dot{+} t_1))m &\leq (q_2^\diamond . t_2)m \\ &= q_2^\diamond (t_2 m) \in \Psi \end{aligned}$$

Hence $(q_2^\diamond . (q_1 \dot{+} t_1))m \in \Psi$. *Since*

$$\begin{aligned} (q_2^\diamond . q_1)m &\leq (q_2^\diamond . q_1)m \oplus (q_2^\diamond . t_1)m = (q_2^\diamond . q_1 \dot{+} q_2^\diamond . t_1)m \\ &= (q_2^\diamond . (q_1 \dot{+} t_1))m \in \Psi, \end{aligned}$$

we have $q_1(q_2^\diamond m) = (q_2^\diamond . q_1)m \in \Psi$. *Now, since* $q_1 \notin (\Psi : V)$ *and* Ψ *is a prime* A*-ideal of* V, *we have* $q_2^\diamond m \in \Psi$. *It follows that* $q_2 m \in \Psi$ *that is a contradiction. Therefore,* $q_1 = q_2$ *and therefore,* $(\Psi : V)$ *is a* Q*-ideal of* A.

Theorem 2. *Let* A *be unital,* V *be a unitary torsion free* A*-semimodule, and* Ψ *be an* A*-ideal of* V. *Then* Ψ *is a prime* $\frac{A}{Ann_A(V)}$*-ideal of* V *if and only if* Ψ *is a prime* A*-ideal of* V.

Proof. Consider Ψ as a prime A-ideal of V. We show that V is an $\frac{A}{Ann_A(V)}$-semimodule. Since $Ann_A(V)$ is a Q-ideal of A, $\frac{A}{Ann_A(V)}$ can be defined. Now, consider operation $\frac{A}{Ann_A(V)} \times V \longrightarrow V$ by $(a + Ann_A(V))x = ax$, for every $a \in A$ and $x \in V$. Then the proof of $(SMV1)$ and $(SMV2)$ are routine.
The proof of $(SMV2)$: Consider $a, b \in Q$. Hence $(a \dot{+} Ann_A(V)) \boxplus (b \dot{+} Ann_A(V)) = c \dot{+} Ann_A(V)$, where $c \in Q$ and $a \dot{+} b \dot{+} Ann_A(V) \subseteq c \dot{+} Ann_A(V)$. Since $(a \dot{+} b \dot{+} Ann_A(V)) \cap (c \dot{+} Ann_A(V)) \neq \emptyset$, we have $a \dot{+} b = c$ and so,

$$\begin{aligned} ((a \dot{+} Ann_A(V)) \boxplus (b \dot{+} Ann_A(V)))x &= (c \dot{+} Ann_A(V))x \\ &= (a \dot{+} b)x \\ &= ax \oplus bx \end{aligned}$$

On the other hand,

$$ax \oplus bx = (a \dot{+} Ann_A(V))x \oplus (b \dot{+} Ann_A(V))x.$$

Then $(SMV2)$ is true.
Let $(a \dot{+} Ann_A(V))x \in \Psi$, for any $a \in Q$ and $x \in V$. Then easily we have $x \in \Psi$ or $(a \dot{+} Ann_A(V)) \in (\Psi : V)$.

Therefore, Ψ is a prime $\frac{A}{Ann_A(V)}$-ideal of V.
Proving reverse mode is easy. □

Theorem 3. *Let A be unital, V be a unitary A-semimodule and Ψ be a proper A-ideal of V. If $a^{\diamond}m \in N$ implies $am \in \Psi$, for every $a \in A$ and $m \in V$, then $(\Psi : V)$ is a prime ideal of A and $\frac{V}{\Psi}$ is a torsion free $\frac{A}{(\Psi : V)}$-semimodule if and only if Ψ is a prime A-ideal of V*

Proof. If $(\Psi : V)$ is a prime ideal of A and $\frac{V}{\Psi}$ is a torsion free $\frac{A}{(\Psi : V)}$-semimodule, then it is routine to show that Ψ is a prime A-ideal of V. Now, let Ψ be a prime A-ideal of V. Since $(\Psi : V)$ is a Q-ideal of A, $\frac{A}{(\Psi : V)}$ is defined. Consider operation $\frac{A}{(\Psi : V)} \times \frac{V}{\Psi} \longrightarrow \frac{V}{\Psi}$ by $(a + (\Psi : V))\frac{x}{\Psi} = \frac{ax}{\Psi}$, for every $a \in A$ and $x \in V$. Then we will have $\frac{V}{\Psi}$ is a torsion free $\frac{A}{(\Psi : V)}$-semimodule. □

IV. Conclusion

In [8], we introduced the new definition of MV-modules. After that, we used that definition to work on modular structures. In this paper, we proved a number of theorems that could be proved by the old definitions only under certain conditions. We intend to expand the use of this definition.

References

[1] R. Cignoli, I. M. L. $D^{'}$Ottaviano and D. Mundici, " Algebraic foundations of many-valued reasoning", Kluwer Academic, Dordrecht, 2000.

[2] A. Di Nola and C. Russo, " Semiring and semimodules issues in MV-algebras", Communications in Algebra, vol. 41, pp. 1017–1048, 2013.

[3] V. Gupta and N. Chaudhari, "Some results on semirings", Radovi Mathematiki, vol. 12, pp. 13–18, 2003.

[4] V. Gupta and N. Chaudhari, "Right π-regular semirings", Sarajevo Journal of Mathematics, vol. 14, pp. 3–9, 2006.

[5] S. Saidi Goraghani and R. A. Borzooei, "On injective MV-modules", Bulletin of of the Section of Logic, vol. 47, pp. 283–298, 2018.

[6] S. Saidi Goraghani and R. A. Borzooei, "Results on prime ideals in PMV-algebras and MV-modules", Italian Journal of Pure and Applied Mathematics, vol. 37, pp. 183–196, 2017.

[7] S. Saidi Goraghani and R. A. Borzooei, "Most results on A-ideals in MV-moduls", Journal of Algebraic Systems, vol. 5, pp. 1–13, 2017.

[8] S. Saidi Goraghani, R. A. Borzooei and S. S. Ahn, "New definition of MV-semimodules", Journal of Intelligent and Fuzzy System, vol. 41, pp. 3715–3726, 2021.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Object Recognition based on Graph theory and Redundant Keypoint Elimination Method

Zahra Hossein-Nejad[1], Mehdi Nasri[2]

[1]*Department of Electrical Engineering, Sirjan Branch, Islamic Azad University, Sirjan, Iran,* hoseinnejad.zahra@yahoo.com
[2]*Department of Biomedical Engineering, Khomeinishahr Branch Islamic Azad University, Isfahan, Iran,* nasri_me@iaukhsh.ac.ir

***Abstract*—Object Recognition System is widely used in different real-life applications such as content-based image retrieval, object detection, etc. In this article, we suggest a novel technique for object detection using Redundant Keypoint Elimination method SIFT-Graph Transformation Matching (RKEMSIFT-GTM). This proposed approach deletes redundant points and eliminates false matches. The proposed improved region-growing, which is a powerful method, is used for the final detection stages. The suggested approach is evaluated on datasets such as COIL-100 and obtained a good recognition rate compared to other detection methods.**

***Index Terms* —Object Recognition; matching; SIFT; GTM.**

I- INTRODUCTION

Object recognition is widely used in the machine vision industry for the aims of inspection, and manipulation[1].Object recognition means the identification of one object seen in the image among a set of defined images or video arrays. The approaches to object recognition could generally be categorized into two: the specific case and the generic case[2]. In the specific case, the aim is to identify the particular object, place or person, such as the Eiffel Tower, my car, and so on. In the generic cases, the aim is to identify objects, or things belonging to the same class, for example, identifying buildings, cars, etc. In these methods, supervised classification methods are used to detect objects. In the first class, the feature-based method can be used to identify specific objects, where matching process is one of the most critical steps. In this method, the accuracy, and quality of object recognition depend on the matching process [3, 4].

Image matching process creates a correspondence between two images of a scene [5] that feature-based matching can be generally divided into two classes, descriptors based matching and special relationships-based matching[6, 7]. In the descriptors method, a descriptor is created using the specifications of the area around each feature. The matches whose distance is more than the threshold value are considered false matches. Matching based on nearest neighbor distance ratio [8] and dual matching [9] are examples. Descriptors-based methods are easily implemented. However, when there are repetitive templates, it leads to the distinctiveness of descriptors and may result in failures in the matching process. The spatial relationships method uses a geometric model among features, geometric model parameters are estimated, and then by calculating the distance between features and spatial information, false matches are removed. This method is used when the detected features are ambiguous, or their neighbors are locally distorted [6]. The GTM [10] and Random Sample Consensus (RANSAC)[11] are examples of spatial relationships based methods.

The RANSAC is a robust approach to delete false matches[12]. This method is not suitable for images with repetitive patterns, and its lack of function is high when the number of false matches is high[13, 14].To overcome these disadvantages, the GTM algorithm can be used.

The GTM was originally introduced by Aguilar [15]. This method uses spatial relationships of the matching points for the false matches deletion [16]. In reference [17], for non-rigid image registration, improved GTM. In reference [18], to improvement mammogram registration images, GTM algorithm is used. In reference [19] to increase correct matches in retinal mosaicing, graph algorithm is used.

Combining descriptors-based matching and GTM increases object recognition precision. SIFT is one of the descriptor-based matching algorithms independent of geometric variations. These advantages have made this algorithm significant to object recognition [20, 21], image registration [22], image fusion [23], and object tracking [24]. Despite the high performance of SIFT algorithm, this algorithm extracts many redundant points that lead to reduction of matching precision [25]. Recently, to solve the mentioned

problem, RKEM-SIFT was presented [26]. In the first step of the algorithm, SIFT keypoints are extracted. Afterwards, distances between different points are computed. When a calculated distance is less than a certain threshold, the redundant keypoint is deleted, and the more important one is kept for the matching. Since redundant keypoints reduce matching accuracy, finally interfere in object recognition.

In this paper, to improve object recognition process, RKEM-SIFT algorithm is used to match the features of objects. Afterwards, the GTM algorithm is used to increase the correct matches and matching precision. Another innovation of the paper is using the region-growing algorithm for detecting the objects boundary. This algorithm is used to identify the boundaries of objects. For this aim, all matches of the GTM are used as seed points of the region-growing algorithm. This selection leads to increase the accuracy of objects' boundary detection in objects with repetitive patterns.

The rest of article as follows. In the section II, releted work on object recognition are presented. In section III, method is presented. In Section IV, evaluation criteria and the test results reviewed. The article is concluded in section V.

II-RELETED WORK ON OBJECT RECOGNITION

In reference[27] , the combination of SIFT, SURF and ORB is used to describe the features and the Random Forest algorithm is used to classify in object recognition. This proposed system works well. In reference[28], SIFT and DRLBP descriptors are used to identify objects. In reference [29], a new technique for pattern recognition is presented, which uses the SIFT for matching and the Hough transform to eliminate false matches. In reference [30], SURF algorithm is used to match and RANSAC is used to eliminate false matches in object recognition. In reference[31], the MAC-RANSAC is proposed to delete false matches in the recognition of multiple objects. This method is more effective than the classic RANSAC. In reference [32], the SURF is used to identify and match features in object recognition. In reference[33] introduces a new system for recognizing objects. In this method, the DBSCAN method is used to calculate the RANSAC parameters, which has increased the accuracy of the object recognition rate. In reference [34], SIFT and morphology operators are used to identify moving objects. In reference [35], the approach for recognizing satellite images is presented, which first uses SIFT to identify features and CNN to classify and identify objects.

III-METHOD

The matching process is one of the significant stages in object recognition [36, 37]. In this section, to improve object recognition, the features and initial matching is carried out using RKEM-SIFT. Then, the secondary matching is done using the GTM. Finally, the object detection is done using the proposed region-growing algorithm according to Fig.1.

A. RKEM SIFT

RKEM-SIFT is an extend version of SIFT, which is used to extract and match keypoints of images. Details of this method are provided in Fig. 2 and 3.

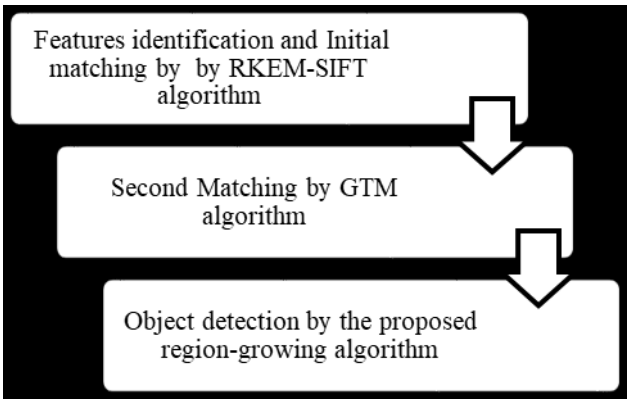


Fig. 1. Flowchart of the proposed object recognition process

1- Generation of scale space

2- Extraction of extreme features

3- Determining the position of the extreme features

4- Creating directions for each keypoint

5- Elimination of redundant keypoints:

- Calculate the distance of keypoints in each image and create a matrix
- Calculate the redundancy index of keypoints
- Determine the threshold value
- The distance of keypoints that are less than the threshold value as redundant points
- The redundant points that have a lower redundancy index are removed

6- Generate keypoints descriptors

7- Initial matching of feature descriptors by NNDR

Fig.2. RKEM-SIFT feature matching diagram

A. Second matching by GTM algorithm

The GTM algorithm based on graph theory is used to delete false matches. Details of this method are provided in an Algorithm 1.

Algorithm: *The GTM Algorithm*
Input: *the set of initial matches*
Output: *Eliminate false matches by GTM*

Calculate the middle distances of the matching points and the middle of the graph matrix in the target object and query object
If the middle matrix of the graph in the target object is not equal to the query object

- *Remove false matches from the middle of the graph matrix in the target object and the query object*
- *Compute middle graph matrix in the target object and query object*

End if

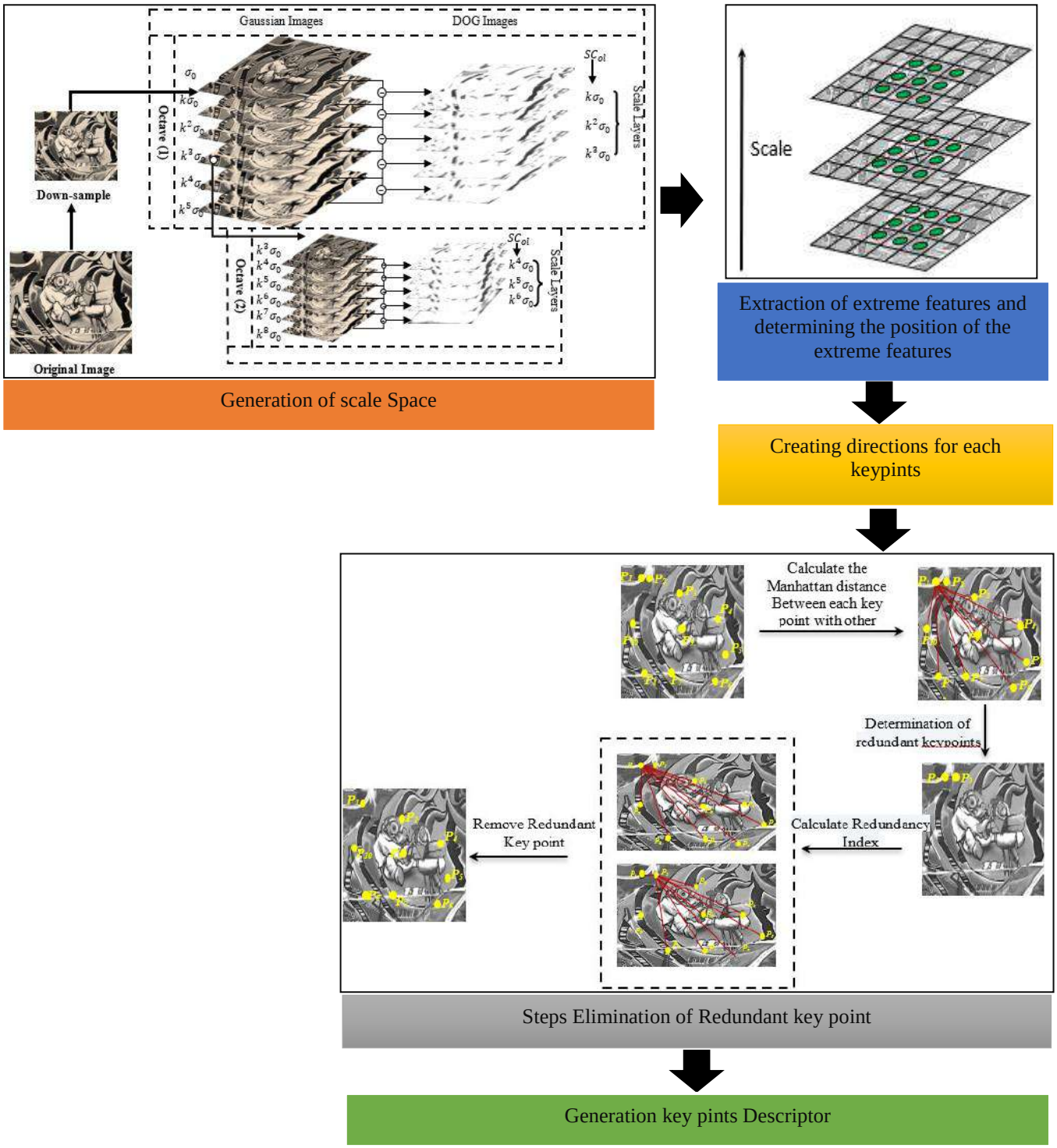


Fig. 3. RKEM-SIFT method in elimination of redundant Keypoints

B. Object detection by the proposed region-growing algorithm

In the proposed approach, the region-growing method is applied on the GTM's matches as seed points. To reach the boundaries of the final image, union operator is applied on detected regions. Final object is considered based on the detected boundaries. This proposed method for identifying objects has a better speed than optimization methods such as CNN. This method cannot be compared with other methods such as optimization methods and threshold-based methods.

IV-IMPLEMENTATION AND EXAMINATION OF RESULTS

Two sets of tests were performed to assess the function of the suggested method in comparison with other recognition methods such as SURF-RANSAC[30], Shi-Tomasi corner detector-SIFT-SURF[38]. For implementation, the software MATLAB®2015A has been used, and all the stages of the suggested method are programmed according to the details stated in section III.

A. Dataset

The two databases used to evaluate the proposed performance are described below.

The COIL-100 data contains 100 objects with angle changes. From each object, are 72 images with different angles which can be seen in Fig.4 examples. Another database that is taken from the Internet and includes ten couples of household objects that can be seen in Fig.5 examples.

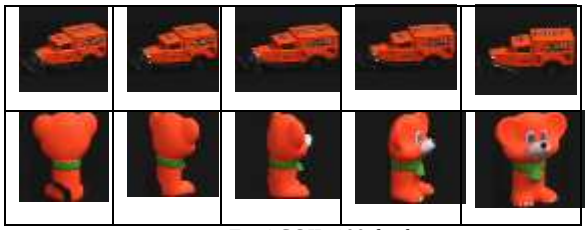


Fig.4.COIL-100 database

Fig. 5. Household objects database

B. Evaluation criteria

The evaluation criteria used in this section are precision (1), correct rate (2) and error rate (3).

$$\text{precision} = \frac{N_c}{N_t} \tag{1}$$

$$\text{correct rate} = \frac{T_c}{N_{object}} \times 100 \tag{2}$$

$$\text{error rate} = \frac{T_f}{N_{object}} \times 100 \tag{3}$$

In (1-3), N_c is the number of correct matches, N_t is the total number of matching, N_{object} is the number of target objects, T_c is the total number of objects correctly detected in the image and T_f is the total number of false objects detected in the image.

C. Examining the suggested approach on the household objects

In this part, ten pairs of natural images are used to check the performance of the object recognition process. Fig.6 shows two examples of object recognition results and Table 1 shows the mean of the results on the ten images.

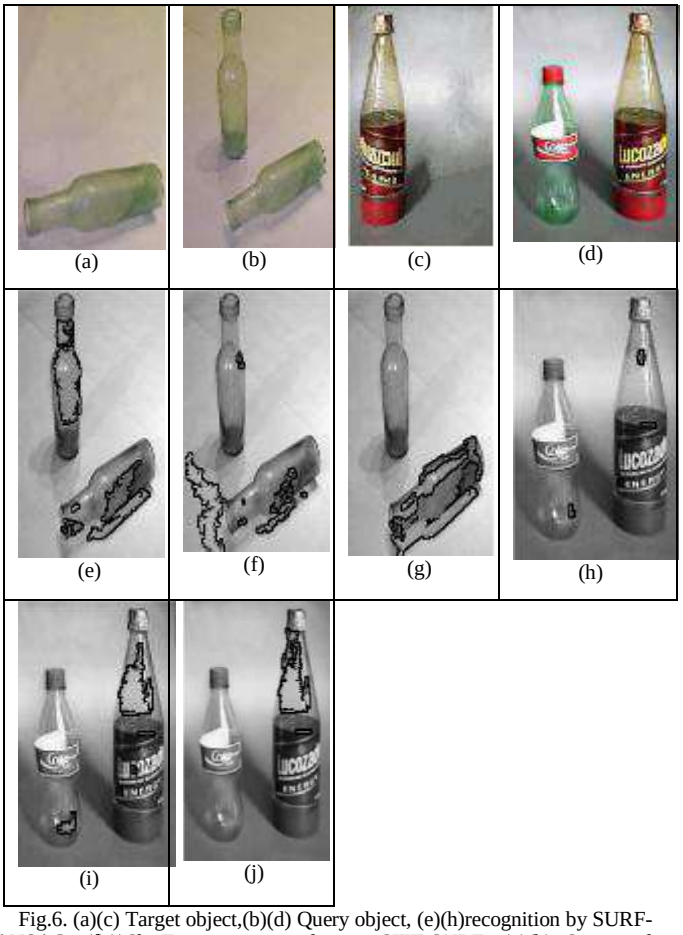


Fig.6. (a)(c) Target object,(b)(d) Query object, (e)(h)recognition by SURF-RANSAC, (f)(i)Shi-Tomasi corner detector-SIFT-SURF, (g)(h) Suggested method

The function of the suggested approach is better than the other methods (SURF-RANSAC, Shi-Tomasi corner detector-SIFT-SURF) in recognition because the suggested method was able to identify the target object well, while other methods (SURF-RANSAC, Shi-Tomasi corner detector-SIFT-SURF) in addition to the target object, other objects were incorrectly identified (Fig.6).

TABLE.I. Recognition results

Method	Precision	Correct rate (%)	error-rate (%)
SURF-RANSAC[30]	0.648	76.13	45.6
Shi-Tomasi corner detector-SIFT-SURF[38]	0.704	79.0	43.2
Suggested proposed	**0.821**	**87.5**	**18.8**

The function of the suggested approach is very effective in correct detection rate, error rate and accuracy compared to other methods of recognition.

D. Examining the suggested approach on the COIL-100 images

In this part, five objects from each object, eleven images with different angles are used to check for recognition (Table II).

TABLE. II. Recognition results on COIL-100

Type image	methods	Correct rate (%)
 Object 1	SURF-RANSAC	90
	Shi-Tomasi corner detector-SIFT-SURF	90
	suggested method	90
 Object 2	SURF-RANSAC	80
	Shi-Tomasi corner detector-SIFT-SURF	70
	suggested method	90
 Object 3	SURF-RANSAC	90
	Shi-Tomasi corner detector-SIFT-SURF	90
	suggested method	100
 Object 4	SURF-RANSAC	70
	Shi-Tomasi corner detector-SIFT-SURF	80
	suggested method	90
 Object 5	SURF-RANSAC	80
	Shi-Tomasi corner detector-SIFT-SURF	80
	suggested method	90

The function of the suggested method is much more appropriate than other methods for identifying complex objects (object 2, object 3, object 4). The proposed method for identifying simple objects (object 1) has almost the same performance as other methods (Table II).

V-CONCLUSION

In the article, a new object recognition algorithm is proposed based on RKEM-SIFT algorithm, GTM, and suggested region-growing segmentation method. At first, important keypoints are extracted based on RKEM-SIFT which removes redundant SIFT keypoints, and keep more important ones. Afterward, GTM algorithm which is based on graph theory is used to delete false matches. The suggested region-growing is the final step to construct the detected object's boundary. Simulation results showed that the suggested approach could detect objects with high accuracy in standard databases such as COIL-100.

REFERENCES

[١] D. G. Lowe, "Object recognition from local scale-invariant features," in *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, 1999, pp. 1150-1157.

[٢] K. Grauman and B. Leibe, "Visual object recognition," *Synthesis lectures on artificial intelligence and machine learning*, vol. 5, pp. 1-181, 2011.

[٣] F. Alhwarin, C. Wang, D. Ristić-Durrant, and A. Gräser, "Improved SIFT-features matching for object recognition," in *Visions of Computer Science-BCS International Academic Conference*, 2008, pp. 179-190.

[٤] D. Hagihara, T. Yano, and N. Kimura, "Parameter Optimization of Feature Matching-Based Object Recognition Algorithm by using CMA-ES," in *2020 16th International Conference on Control, Automation, Robotics and Vision*) *ICARCV)*, 2020, pp. 866-873.

[٥] Z. Hossein-Nejad and M. Nasri, "Clustered redundant keypoint elimination method for image mosaicing using a new Gaussian-weighted blending algorithm," *The Visual Computer*, 2021/07/19 2021.

[٦] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and vision computing*, vol. 21, pp. 977-1000, 2003.

[٧] J. Liu, G. Zeng, and J. Fan, "Fast local self-similarity for describing interest regions," *Pattern Recognition Letters*, vol. 33, pp. 1224-1235, 2012.

[٨] D. G .Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, pp. 91-110, 2004.

[٩] S. Wang, H. You, and K. Fu, "BFSIFT: A novel method to find feature matches for SAR image registration," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, pp. 649-653, 2012.

[١٠] W. Aguilar, Y. Frauel, F. Escolano, M. E. Martinez-Perez, A. Espinosa-Romero, and M. A. Lozano, "A robust graph transformation matching for non-rigid registration," *Image and Vision Computing*, vol. 27, pp. 897-910, 2009.

[١١] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, pp. 381-395, 1981.

[١٢] Z. Hossein-nejad and M. Nasri, "Image registration based on SIFT features and adaptive RANSAC transform," in *Communication and Signal Processing (ICCSP), 2016 International Conference on*, 2016, pp. 1087-1091.

[١٣] Z. Hossein-Nejad and M. Nasri, "A-RANSAC: Adaptive random sample consensus method in multimodal retinal image registration," *Biomedical Signal Processing and Control*, vol. 45, pp. 325-338, 2018.

[١٤] A. Hast, J. Nysjö, and A. Marchetti, "Optimal ransac-towards a repeatable algorithm for finding the optimal set," 2013.

[١٥] W. Aguilar, Y. Frauel, F. Escolano, M. E. Martinez-Perez, A. Espinosa-Romero, and M. A. Lozano, "A robust graph transformation matching for non-rigid registration," *Image Vision Computing*, vol. 27, pp. 897-910, 2009.

[١٦] W. Aguilar, "Object recognition based on the structural correspondence of local features," *MsThesis, UNAM, México*, 2006.

[١٧] M. Izadi and P. Saeedi, "Robust weighted graph transformation matching for rigid and nonrigid image registration," *IEEE Transactions on Image Processing*, vol. 21, pp. 4369-4382, 2012.

[١٨] Y.-j. Zhong and S.-F. Chen, "An Improved Algorithm Based on SIFT and Graph Transformation for Mammogram Registration," *Journal of Convergence Information Technology*, vol. 7, pp. 279-285, 2012.

[١٩] W. Aguilar, M. E. Martinez-Perez, Y. Frauel, F. Escolano, M. A. Lozano, and A. Espinosa-Romero, "Graph-based methods for retinal mosaicing and vascular characterization," in *International Workshop on Graph-Based Representations in Pattern Recognition*, 2007, pp.٣٦-٢٥ .

[٢٠] L. Lenc and P. Král, "Automatic face recognition system based on the SIFT features," *Computers & Electrical Engineering*, vol. 46, pp. 256-272, 2015.

[٢١] S. A. M. Gilani, "Object recognition by modified scale invariant feature transform," in *Semantic Media Adaptation and Personalization, 2008. SMAP'08. Third International Workshop on*, 2008, pp. 33-39.

[٢٢] H.-H. Chang and W.-C. Chan, "Automatic Registration of Remote Sensing Images Based on Revised SIFT With Trilateral Computation and Homogeneity Enforcement," *IEEE Transactions on Geoscience and Remote Sensing*, 2021.

[٢٣] Y. Jiao, "Optimization and Innovation of SIFT Feature Matching Algorithm in Static Image Stitching Scheme," in *Journal of Physics: Conference Series*, 2021, p. 032017.

[٢٤] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," *Acm computing surveys (CSUR)*, vol. 38, pp. 1-45, 2006.

[٢٥] Z. Hossein-Nejad and M. Nasri, "Retinal Image Registration based on Auto-Adaptive SIFT and Redundant Keypoint Elimination Method," in *2019 27th Iranian Conference on Electrical Engineering (ICEE)*, 2019, pp. 1294-1297.

[٢٦] Z. Hossein-Nejad and M. Nasri, "RKEM: Redundant Keypoint Elimination Method in Image Registration," *IET Image Processing*, 2017.

[٢٧] M. Bansal, M. Kumar, and M. Kumar, "2D object recognition: a comparative analysis of SIFT, SURF and ORB feature descriptors," *Multimedia Tools and Applications*, vol. 80, pp. 18839-18857, 2021.

[٢٨] M. S. A. Atole and S. Khomane, "OBJECT RECOGNITION USING COMBINED DRLBP AND SIFT FEATURES," 2020.

[٢٩] I. Nurhaida, A. Noviyanto, R. Manurung, and A. M. Arymurthy, "Automatic Indonesian's batik pattern recognition using SIFT approach," *Procedia Computer Science*, vol. 59, pp. 567-576, 2015.

[٣٠] S. Kole, C. Agarwal, T. Gupta, and S. Singh, "Surf and ransac: A conglomerative approach to object recognition," *International Journal of Computer Applications*, vol. 109, 2015.

[٣١] J. Rabin, J. Delon, Y. Gousseau, and L. Moisan, "MAC-RANSAC: a robust algorithm for the recognition of multiple objects ",in *Fifth International Symposium on 3D Data Processing, Visualization and Transmission (3DPTV 2010)*, 2010, p. 051.

[٣٢] B. Bhosale Swapnali, S. Kayastha Vijay, and K. HarpaleVarsha, "Feature extraction using surf algorithm for object recognition," *International journal of technical research and applications*, vol. 2, pp. 197-199, 2014.

[٣٣] R. Scitovski, S. Majstorović, and K. Sabo, "A combination of RANSAC and DBSCAN methods for solving the multiple geometrical object detection problem," *Journal of Global Optimization*, vol. 79, pp. 669-686, 2021.

[٣٤] C. Zhou, J. Xing, and H. Liu, "Multiple Properties-Based Moving Object Detection Algorithm," *Journal of Information Processing Systems*, vol. 17, pp. 124-135, 2021.

[٣٥] K.-Z. Liu, P.-J. Lee, G.-C. Xu, and B.-H. Chang, "SIFT-Enhanced CNN Based Objects Recognition for Satellite Image," in *2020 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-Taiwan)*, 2020, pp. 1-2.

[٣٦] B. Sirmacek and C. Unsalan, "Urban-area and building detection using SIFT keypoints and graph theory," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, pp. 1156-1167, 2009.

[٣٧] C. Tao, Y. Tan, H. Cai, and J. Tian, "Airport detection from large IKONOS images using clustered SIFT keypoints and region information ",*IEEE Geoscience and Remote Sensing Letters*, vol. 8, pp. 128-132, 2011.

[٣٨] M. Bansal, M. Kumar, M. Kumar, and K. Kumar, "An efficient technique for object recognition using Shi-Tomasi corner detection algorithm," *Soft Computing*, vol. 25, pp. 4423-4432.٢٠٢١ ,

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

On NEUTRO G-SUBALGEBRA

1st Mohammad Hamidi, 2nd Marzieh Rahmati

1st Department of Mathematics, University of Payame Noor, Tehran, Iran, m.hamidi@pnu.ac.ir
2nd Department of Mathematics, University of Payame Noor, Tehran, Iran, m.rahmati@pnu.ac.ir

***Abstract*— The aim of this paper is to present a study about recent progressions in the study of neutrosophic algebraic structures. This paper introduces the novel concept of neutro-G-algebra and shows that neutro-G-algebra are different from G-algebra. More precisely, we present several results from the study of certain properties of neutro-G-algebras.**

***Index Terms*—G-algebra, neutro-G-algebra.**

I. Introduction

Since 1966 many algebraic structures are introduced such as BCK-algebras, BCH-algebra, BCI-algebra and B-algebra. In 2012, R. K. Bandru and N. Rafi introduced a new notion called G-algebra [2]. The notion of G-algebra was a generalization of QS-algebra. It has been shown that every QS-algebra is a G-algebra but the converse need not, in general, be true [2]. This notion plays an important role in algebra and has many applications [1, 5].

The notion of a neutrosophic set was given by Samarandache [8, 10, 11]. An neutrosophic set is the generalization of a crisp set, fuzzy sets and intuitionistic fuzzy set. The membership of truth (TM) and falsehood (FM) and indeterminacy (IM) are independent in neutrosophic set. The role of indeterminacy is very important for decision-making problems. In several areas of applied mathematics such as pattern recognition, machine learning, health care, computer science, and statistics, decision-making problems, medical diagnosis, image processing, educational problems, and dispute resolution, an neutrosophic set and its extensions play a critical role. Neutrosophic Sets and Systems international journal (which is in Scopus and Web of Science) is a tool for publications of advanced studies in neutrosophy, neutrosophic set, neutrosophic logic, neutrosophic probability, neutrosophic statistics, neutrosophic measure, neutrosophic integral, and so on, studies that started in 1995 and their applications in any field, such as the neutrosophic structures developed in algebra, geometry, topology, etc. Recently, Florentin Smarandache generalized the classical Algebraic Structures to NeutroAlgebraic Structures NeutroAlgebras) and AntiAlgebraic Structures (AntiAlgebras) and he proved that the NeutroAlgebra is a generalization of Partial Algebra[9]. Neutrosophy is a new branch of philosophy that studies triads of the form $(< A >, < neutA >, < antiA >)$, where $< A >$ is an entity (i.e., element, concept, idea, theory, logical proposition, etc.), $< antiA >$ is the opposite of $< A >$, while $< neutA >$ is the neutral (or indeterminate) between them, i.e., neither $< A >$ nor $< antiA >$. Based on neutrosophy, the neutrosophic triplets were founded; they have a similar form: $(x, neut(x), anti(x))$, that satisfy some axioms, for each element x in a given set [3, 6, 7]. A NeutroAlgebra is an algebra which has at least one Neutro Operation (that is well-defined (also called inner-defined) for some elements, indeterminate for others, and outer-defined for the others) or one NeutroAxiom (axiom that is true for some elements, indeterminate for other elements, and false for the other elements). Florentin Smarandache Through a theorem proved that NeutroAlgebra is a generalization of Partial Algebra, and examples of NeutroAlgebras that are not partial algebras were given. Also, the NeutroFunction and NeutroOperation were introduced.[9]

Regarding these points, we now introduce the concept of neutro-G-algebras based on axioms of G-algebras, but having a different outcome. In the system of G-algebras, the operation is totally well-defined for any two given elements, but in neutro-G-algebras its outcome may be well-defined, outer-defined, or indeterminate. Any G-algebra is a system which considers that all its axioms are true; but we weaken the conditions that the axioms are not necessarily totally true, but also partially false, and partially indeterminate. So, one of our main motivation is a weak coverage of the classical axioms of G-algebras.

II. neutro G-subalgebra

Let $\mathbb{N}$ be a set of natural numbers and $n \in \mathbb{N}$ and X, Y be two non-empty sets. Then an n-ary operation $\circ : X^n \to Y$ is called a NeutroOperation if it has $x \in X^n$ for which $\circ(x)$ is well-defined (degree of truth (T)), $x \in X^n$ for which $\circ(x)$ is indeterminate (degree of indeterminacy (I)), and $x \in X^n$ for which $\circ(x)$ is outer-defined (degree of falsehood (F)), where $T, I, F \in [0, 1]$, with $(T, I, F) \neq (1, 0, 0)$ that represents the n-ary (total, or classical) Operation, and $(T, I, F) \neq (0, 0, 1)$ that represents the n-ary AntiOperation. Again, in this definition "neutro" stands for neutrosophic, which means the

existence of outer-ness, or undefined-ness, or unknown-ness, or indeterminacy in general. In this regards, for any given set X, we classify n-ary operation on X^n by (i); (**classical**) **Operation** is an operation well-defined for all set's elements, (ii); **NeutroOperation** is an operation partially well-defined, partially indeterminate, and partially outer-defined on the given set and (iii); **AntiOperation** is an operation outer-defined for all set's elements.

Definition 1. *[4] Let X be a non-empty set with a binary operation " $*$ " and a constant "0". Then, $(X, *, 0)$ is called a G-algebra if it satisfies the following conditions:*

$(BCI\text{-}1)$ $x * x = 0$,
$(BCI\text{-}2)$ $x * (x * y) = y$, *for all* $x, y \in X$.

Definition 2. *Let X be a non-empty set, $0 \in X$ be a constant and " $*$ " be a binary operation on X. An algebra $(X, *, 0)$ of type $(2, 0)$ is said to be a neutro-G-algebra, if it satisfies the following NeutroAxioms:*
$(NBCI\text{-}1)$ $\big(\exists\ x \in X$ *such that* $x * x = 0\big)$ *and* $\big(\exists\ x \in X$ *such that* $x * x \neq 0$ *or indeterminate*$\big)$;
$(NBCI\text{-}2)$ $\big(\exists\ x, y \in X$ *such that* $x * (x * y) = y\big)$ *and* $\big(\exists\ x, y \in X$ *such that* $x * (x * y) \neq y$ *or indeterminate*$\big)$;
Each above NeutroAxiom has a degree of equality (T), degree of non-equality (F), and degree of indeterminacy (I), where $(T, I, F) \notin (1, 0, 0), (0, 0, 1)$.

Example 1. *Let $X = \{0, 1, 2, 3\}$ be a set with the following table:*

$*$	0	1	2	3
0	0	0	0	0
1	1	0	0	0
2	2	0	0	0
3	3	3	3	0

Then $(X, *, 0)$ there is no G-algebra because $(BCI\text{-}2)$ does not apply and there is no neutro-G-algebra because $(NBCI\text{-}1)$ does not apply.

Example 2. *Let $X = \mathbb{Z}$. Then $(X, *, 1)$ is a neutro-G-algebra, where for all $x, y \in X$, we have $x * y = xy$.*

Let $X \neq \emptyset$ be a finite set. We denote $\mathcal{N}_G(X)$ by the set of all neutro-G-algebras that is constructed on X, respectively.

Theorem 1. *Every G-algebra, can be extended to a Neutro-G-algebra*

Proof: Let $(X, *, 0)$ be a G-algebra and $\alpha \notin X$, and U be the universe of discourse that strictly includes $X \cup \alpha$. For all $x, y \in X \cup \{\alpha\}$, define $*_\alpha$ on $X \cup \{\alpha\}$ by $x *_\alpha y = x * y$ where $x, y \in X$ and if $\alpha \in \{x, y\}$, define $x *_\alpha y$ as indeterminate or $x *_\alpha y \notin X \cup \alpha$. Then $(X \cup \alpha, *_\alpha, 0)$ is a Neutro-G-algebra. ∎

Example 3. *Let $X = \{0, 1, 2, 3, 4, 5\}$ and Consider Table I. Then*

(i) $(X, *_1, 0)$ *is a neutro-G-algebra and* $(X \setminus \{3, 4, 5\}, *_1, 0)$ *is a G-algebra.*

TABLE I
NEUTRO-G-ALGEBRAS

$*_1$	0	1	2	3	4	5
0	0	1	2	0	3	1
1	1	0	2	4	2	3
2	2	1	0	5	0	4
3	2	4	3	5	0	1
4	0	4	1	1	2	0
5	1	3	5	0	2	a

$*_2$	0	1	2	3	4	5	6	7
0	0	2	1	3	4	5	0	1
1	1	0	3	2	5	0	7	3
2	2	3	0	1	6	7	1	4
3	3	2	1	0	7	6	2	5
4	4	5	6	7	0	3	1	0
5	5	4	7	6	1	0	4	6
6	1	2	3	7	6	0	z	1
7	2	5	6	3	2	4	0	y

(ii) *If $z \in \{1, 2, 3, 5, 7\}$ and $y \in \{1, 2, 3, 4, 5, 6\}$, then $(X \cup \{6, 7\}, *_2, 0)$ is a neutro-G-algebra and $(X, *_2, 0)$ is a G-algebra.*

Remark 1. *We observe that the two axioms $(NBCI\text{-}1)$ and $(NBCI\text{-}2)$ are independent.*
Let $X = \{0, 1, 2\}$ be a set with the following left Table. Then the axiom$(NBCI\text{-}2)$ holds but because does not exist $x \in X$ such that $x *_1 x = 0$, the axiom $(NBCI\text{-}1)$ not established. Similarly, the set $X = \{0, 1, 2\}$ with the following right Table satisfy the axiom $(NBCI\text{-}1)$ holds but because does not exist $x, y \in X$ such that $x *_2 (x *_2 y) \neq y$, the axiom $(NBCI\text{-}2)$ not established. (Consider Table II)

TABLE II
NEUTRO-G-ALGEBRAS

$*_1$	0	1	2
0	2	0	2
1	1	1	0
2	0	1	2

and

$*_2$	0	1	2
0	0	1	2
1	1	0	2
2	0	1	2

.

Proposition 1. *If $(X, *, 0)$ is a neutro-G-algebra, then the following conditions hold:*
$(NBCI\text{-}3)$ $\big(\exists\ x \in X$ *such that* $x * 0 = x\big)$ *and* $\big(\exists\ x \in X$ *such that* $x * 0 \neq x$ *or indeterminate*$\big)$.
$(NBCI\text{-}4)$ $\big(\exists\ x \in X$ *such that* $0 * (0 * x) = x\big)$ *and* $\big(\exists\ x \in X$ *such that* $0 * (0 * x) \neq x$ *or indeterminate*$\big)$.

Proof: $(NBCI\text{-}3)$ Since $(X, *, 0)$ be a Neutro-G-algebra, By definition, $\big(\exists\ x \in X$ such that $x * 0 = x * (x * x) = x\big)$ and $\big(\exists\ x \in X$ such that $x * 0 = x * (x * x) \neq x\big)$.
$(NBCI\text{-}4)$ Put $x = 0$ and $y = x$ in $(NBCI\text{-}2)$, then we get $(NBCI\text{-}4)$. ∎

Proposition 2. *Let $(X, *, 0)$ be a neutro-G-algebra. Then the following conditions hold:*
(i) $\big(\exists\ x, y \in X$ *such that* $(x * (x * y)) * y = 0\big)$ *and* $\big(\exists\ x, y \in X$ *such that* $(x * (x * y)) * y \neq 0$ *or indeterminate* $\big)$;

(ii) $\big(\exists\ x,y \in X$, *such that if* $x*y=0$, *we have* $x=y\big)$ *and* $\big(\exists\ x,y \in X$, *such that if* $x*y=0$, *we have* $x \neq y$ *or indeterminate* $\big)$;

(iii) $\big(\exists\ x,y \in X$, *such that if* $0*x=0*y$, *we have* $x=y\big)$ *and* $\big(\exists\ x,y \in X$, *such that if* $0*x=0*y$, *we have* $x \neq y$ *or indeterminate* $\big)$.

Proof:

(i) By $(NBCI\text{-}2)$, $\big(\exists\ x,y \in X$ such that $(x*(x*y))*y=y*y\big)$. Thus By $(NBCI\text{-}1)$, $\big(\exists\ x,y \in X$ such that $(x*(x*y))*y=y*y=0\big)$ and $\big(\exists\ x,y \in X$ such that $(x*(x*y))*y=y*y\neq 0\big)$.

(ii) Let $x*y=0$, then $x*(x*y)=x*0$. By $(NBCI\text{-}2)$ and $(NBCI\text{-}3)$ $\big(\exists\ x,y \in X$ such that $y=x*(x*y)=x*0=x\big)$ and $\big(\exists\ x,y \in X$ such that $y=x*(x*y)=x*0\neq x\big)$.

(iii) Let $0*x=0*y$, then $0*(0*x)=0*(0*y)$. By $(NBCI\text{-}2)$ $\big(\exists\ x,y \in X$ such that $x=0*(0*x)=0*(0*y)=y\big)$ and $\big(\exists\ x,y \in X$ such that $x\neq 0*(0*x)=0*(0*y)\neq y\big)$. ∎

Theorem 2. *If* $(X,*,0)$ *is a neutro-G-algebra satisfying* $\big(\exists\ x,y \in X$ *such that* $(x*y)*(0*y)=x\big)$ *and* $\big(\exists\ x,y \in X$ *such that* $(x*y)*(0*y)\neq x\big)$. *Then* $\big(\exists\ x,y,z \in X$ *such that* $x*z=y*z$ *implies* $x=y\big)$ *and* $\big(\exists\ x,y,z \in X$ *such that* $x*z=y*z$ *implies* $x\neq y$ *or indeterminate* $\big)$.

Proof: Let $(X,*,0)$ be a neutro-G-algebra satisfying $\big(\exists\ x,y \in X$ such that $(x*y)*(0*y)=x\big)$ and $\big(\exists\ x,y \in X$ such that $(x*y)*(0*y)\neq x\big)$. If there exists $x,y,z \in X$ such that $x*z=y*z$. Thus $(x*z)*(0*z)=(y*z)*(0*z)$ implies $x=y$ and $x\neq y$. ∎

Theorem 3. *Let* $(X,*,0)$ *be a neutro-G-algebra. Then the following are equivalent.*

(i) $\big(\exists\ x,y,z \in X$ *such that* $(x*y)*z=(x*z)*y\big)$ *and* $\big(\exists\ x,y,z \in X$ *such that* $(x*y)*z\neq(x*z)*y$ *or indeterminate* $\big)$.

(ii) $\big(\exists\ x,y,z \in X$ *such that* $(x*y)*(x*z)=z*y\big)$ *and* $\big(\exists\ x,y,z \in X$ *such that* $(x*y)*(x*z)\neq z*y$ *or indeterminate* $\big)$.

Proof: $(i)\Rightarrow(ii)$:Assume (i). Then there exists $x,y,z\in X$ such that $(x*y)*(x*(x*z))=(x*y)*z=(x*z)*y$ and there exists $x,y,z\in X$ such that $(x*y)*(x*(x*z))=(x*y)*z\neq(x*z)*y$.
$(ii)\Rightarrow(i)$:Assume (ii). Then there exists $x,y,z\in X$ such that $(x*y)*(x*z)=z*y=(x*(x*z))*y$ and there exists $x,y,z\in X$ such that $(x*y)*(x*z)\neq z*y=(x*(x*z))*y$. ∎

Lemma 1. *Let* $(X,*,0)$ *be a neutro-G-algebra. Then* $\big(\exists\ a,x,y \in X$ *such that* $a*x=a*y$ *implies* $x=y\big)$ *and* $\big(\exists\ a,x,y \in X$ *such that* $a*x=a*y$ *implies* $x\neq y$ *or indeterminate*$\big)$.

Proof: Let X be a neutro-G-algebra. Since there exists $a,x,y\in X$ such that $a*x=a*y$. We get that $a*(a*x)=a*(a*y)\Rightarrow x=y$ and there exists $a,x,y\in X$ such that $x\neq a*(a*x)=a*(a*y)\neq y$. ∎

Theorem 4. *Let* $(X,*,0)$ *be a neutro-G-algebra. Then the following are equivalent.*

(i) $\big(\exists\ x,y \in X$ *such that* $(x*y)*(x*z)=z*y\big)$ *and* $\big(\exists\ x,y \in X$ *such that* $(x*y)*(x*z)\neq z*y$ *or indeterminate* $\big)$.

(ii) $\big(\exists\ x,y \in X$ *such that* $(x*z)*(y*z)=x*y\big)$ *and* $\big(\exists\ x,y \in X$ *such that* $(x*z)*(y*z)\neq x*y$ *or indeterminate* $\big)$.

Proof: $(i)\Rightarrow(ii)$:Assume (i). There exists $x,y,z\in X$ such that $(x*y)*(x*z)=z*y$. Then $(x*y)*((x*y)*(x*z))=(x*y)*(z*y)$, thus by $(NBCI\text{-}2)$; $(x*z)=(x*y)*(z*y)$. According to the assumption there exists $x,y,z\in X$ such that $(x*y)*(x*z)\neq z*y$. Then $(x*y)*((x*y)*(x*z))\neq(x*y)*(z*y)$, thus by $(NBCI\text{-}2)$; $(x*z)\neq(x*y)*(z*y)$.
$(ii)\Rightarrow(i)$:Assume (ii). There exists $x,y,z\in X$ such that $(x*z)*(y*z)=x*y$, then by $(NBCI\text{-}2)$ and Lemma 1 $(x*z)*(y*z)=(x*z)*((x*z)*(x*y))$ thus $(y*z)=(x*z)*(x*y)$.According to the assumption there exists $x,y,z\in X$ such that $(x*z)*(y*z)\neq x*y$, then by $(NBCI\text{-}2)$ and Lemma 1 $(x*z)*(y*z)\neq(x*z)*((x*z)*(x*y))$ thus $(y*z)\neq(x*z)*(x*y)$. ∎

CONCLUSION

In this paper, we have introduced the concept of neutro-G-algebras and studied their properties. The Neutro-G-algebra has many utilizations in different areas, where we connect Neutro-G-algebra to other sciences such as economics, computer sciences and other engineering sciences. I hope this work would serve as a foundation for further studies on the structure of neutro-G-algebras.

REFERENCES

[1] D. Al-Kadi, R. Hosny, "Notes on G-algebra and its derivations," Math Sci Lett, **6**, 2017, no. 3, pp.287–292.

[2] R. K. Bandru, N. Rafi, "On G-algebras." Scientia Magna International Book Series , **8** ,2012, no.3, pp.1–7.

[3] M. Hamidi, F. Smarandache,"Neutro-BCK-algebra," INS. J. Neutrosophic Sci., **8** ,2020, no. 2, pp. 110–117.

[4] Y. Imai and K. Iseki, "It On axioms systems of propositional calculi." XIV, Proc. Japan Academy, **42**, 1966, pp.19–22.

[5] J. Neggers, HS. Kim, "A fundamental theorem of B-homomorphism for B-algebras." Int Math J, **2** ,2002, no.3, pp.207–214.

[6] A. Rezaei, F. Smarandache, "The Neutrosophic Triplet of BI-algebras," Neutrosophic Sets and Systems, **33** ,2020, pp. 313–321.

[7] A. Rezaei, F. Smarandache, "On Neutro-BE-algebras and Anti-BE-algebras," Int. J. Neutrosophic Sci., **4** ,2020, no. 1, pp. 08–15.

[8] F. Smarandache, "A Unifying Field in Logics: Neutrosophic Logic, Neutrosophy, Neutrosophic Set, Neutrosophic Probability," American Research Press, Rehoboth., 2003.

[9] F. Smarandache, "Neutro algebra is a generalization of partial algebra," Int. J. Neutrosophic Sci., **2**, 2020, no.1, pp.8–17.

[10] F. Smarandache,“Neutrosophy: Neutrosophic probability, set, and logic: Analytic synthesis and synthetic analysis,”American Research Press, 1998.
[11] F. Smarandache,“Neutrosophic set-a generalization of the intuitionistic fuzzy set,” International Journal of Pure and Applied Mathematics, **24**, 2005, no.1, pp.287–297.

9th Iranian Joint Congress on Fuzzy and Intelligent Systems
(CFIS2022)

March 2-4, 2022, Higher Education Complex of Bam

Seven Staged Identity Recognition System Using Kinect V.2 Sensor

Seyed Muhammad Hossein Mousavi [1], Atiye Ilanloo [2]

[1] *Independent Scientist, Tehran, Iran,* mosavi.a.i.buali@gmail.com
[2] *Faculty of Humanities- Psychology, Islamic Azad University-Rasht, Gilan, Iran*

***Abstract*— By employing artificial intelligence techniques and algorithms such as color and depth image processing, signal processing, machine learning, evolutionary algorithms and fuzzy systems, an identity recognition expert system with approximate recognition accuracy of 99% is proposed. Available identity recognition systems mostly are in three stages which may lead to some security problems, so it is decided to make a robust system. Proposed system uses Kinect Version 2 sensor in order to conduct 7 main stages of recognition. The system includes following stages of recognition and estimation which are, face and voice recognition, finger print recognition, iris recognition, gesture recognition, sex detection and age estimation. By adding macro lens to the sensor, recognition accuracy for fingerprint and iris increases significantly. All efforts on this project were to achieve the highest potential out of available techniques. The system is learning based and has high precision and could be well used in industrial purposes. By installing macro lens on Kinect sensor, the system could compete with other expensive identification systems. It has to be mention that proposed system works well in the pure darkness ass Kinect sensor supports the infrared spectrum.**

***Index Terms* — Expert system; Security; Kinect sensor; Identity recognition; Image and signal processing; Macro lens**

I. INTRODUCTION

As technology enhances, possibility of using full potential of it gets higher. A lot of organizations and even personal sites need to have a proper security system which mostly uses image processing devices like cameras. Also, most of the use just color sensor, not depth. But it could be developed to a higher degree when a security breach could occur. Obviously using more security technologies and devices helps to the rising problem. As it mentioned, available identity recognition systems are mostly in three stages. It is mostly enough, but making more robust system is always sensible. Proposed system uses Kinect sensor version 2 [1] and high level of artificial intelligence algorithms and techniques such as image and signal processing [2], data mining [3], machine learning [4], fuzzy sets [5] and evolutionary algorithms [6] in order to make the final results as better as possible. System performs voice recognition [7], face recognition [8], fingerprint recognition [9], iris recognition [10], gesture recognition [11], sex detection and age estimation [12] tasks. Also, a macro lens is installed on the color sensor in order to increase the receiving details. In the other hand depth sensor of the Kinect, overcomes the different lighting environment for the final use.

II. PRIOR RELATED RESEARCHES

Akihiro Machida in 2006, succeeded to invent a touchless fingerprint recognition system which was assisted by a remote control under the patent number of US7116805B2 [13].

In the field of sex detection and age estimation Yamazaki, Kazuhiro, et al, could invent a system which employed spatial and frequency domains features to perform the task. This patent is registered under the US10204266B2 serial in 2011 [14].

Another age estimation method which uses Kinect sensor is belong to SMH, Mousavi in 2018. System is real time based and so fast which works by just a single image even in pure darkness [12].

Anter Abozaid and et al, made a system for face and voice recognition purposes in 2018. They used eigenface and Principal Component Analysis (PCA) features and classified their systems by Gaussian Mixture Model (GMM) [15], Artificial Neural Network (ANN), and Support Vector Machine (SVM) [8].

Another research belong to face recognition is for Mousavi, Seyed Muhammad Hossein, and S. Younes Mirinezhad in 2021, which they used Convolutional Neural Networks [16] to train their system [17].

Peng, Chang, et al, made an ultrasonic fingerprint recognition system in 2021 for the first time. Basically, they used signal processing instead of image [9].

For iris recognition task, Muthazhagan, Balaji, and Suriya Sundaramoorthy employed a deep learning algorithm called CNN on their systems, achieving high recognition accuracy [10].

Fenglin Liu, et al succeed to make a gesture recognition system using Kinect sensor and RGB network with high accuracy in 2019 [11].

III. PROPOSED IDENTITY RECOGNITION METHOD

A. Advantages

The system has 7 stages of estimation, detection and recognition by color and depth data using Kinect V.2 Sensor. As the system employs infrared sensor, some stages could perform in pure darkness. It is an expert system, means that it could be replaced for human and never gets tired or error. It is cheaper than its similar types as it is using cheap sensor. Fingerprint recognition is touchless and does not need to professional tool for keeping it. Also, macro lens increases the final recognition accuracy for iris and fingerprint, significantly. As depth data is known as 2.5-Dimensional (2.5-D), it could be easily converted to 3-Dimensional (3-D) model, which is an advantage. If one of the steps does not work with enough accuracy, system get into halt situation. Also, it has ease of use for final user and has pretty convenient speed.

B. Usage

This expert system could be used in security and criminology systems, law schools, defense industry research, museum preservation, museum security of banknotes and ingots in banks, storage of the latest available technologies that are not available to the public, storage of the above documents Confidential, secret and top secret that should not be disclosed, where the latest achievements of medical science are kept, such as deadly viruses that should not be taken out of the building, and so on. In general, in any place where the building configuration is solid, has valuable objects or documents, and the number of visits and people visiting is low.

C. Steps

System starts with acquiring color and depth data from Kinect sensor and system announces “Please stand in front of the sensor (1 meter)”. Second step is belonged to gesture, voice, face recognitions, sex detection and age estimation. In gesture recognition step, body will be extracted from depth data and Gabor features [17, 21] will be extracted in the frequency domain [18, 21] from the data. Voice recognition is consisted of noise removal and signal normalization by median filter followed by extracting Mel-Frequency Cepstral Coefficients (MFCCs) features [2] in the frequency domain. Face recognition however has some preprocessing steps of background removal and face extraction by viola and jones algorithm [19] in color and depth data followed by applying multi directional edge detection (canny) [20] and closing morphological operation plus edge sharpening by unsharp masking [20]. Then the data sends for CNN for final classification. Age estimation and sex detection steps have the same preprocessing stage of face recognition step, but for age estimation Local Phase Quantization LPQ [21] features are extracted in frequency domain from color and depth data and normalizes in the range of 0 to 1. Also, for sex detection step, Speeded Up Robust Features SURF [21] are extracted from color data in spatial domain. Now, all extracted features from all steps goes for dimensionality reduction step conducted by Lasso regularization algorithm [21] for removing outliers and faster classification speed. Here in classification step final data of gestures and sex goes for classification by Deferential Evolution-ANFIS [22, 24] classification algorithm and other steps simply use SVM. Now it is time for fingerprint and iris recognition. System announces “Please hold your finger in 3–4-centimeter distance from sensor” and data acquisition in color and depth starts. For fingerprint recognition, CNN is used and for iris recognition another announcement “Please hold your right eye in 3–4-centimeter distance from the sensor”, system says. Afterwards, Histogram of Oriented Gradient (HOG) [17, 21] features got extracted from color data followed by shallow neural network classification algorithm. Finally expert system has to check if all stages done successfully and matches the database or not. If yes, identity recognition completed and electronic lock gets unlocked followed by halting the system. If not, system returns to stage 1 for reidentification. Figure 1 represents experiment environment and setup. Figure 2 shows the flowchart of proposed method.

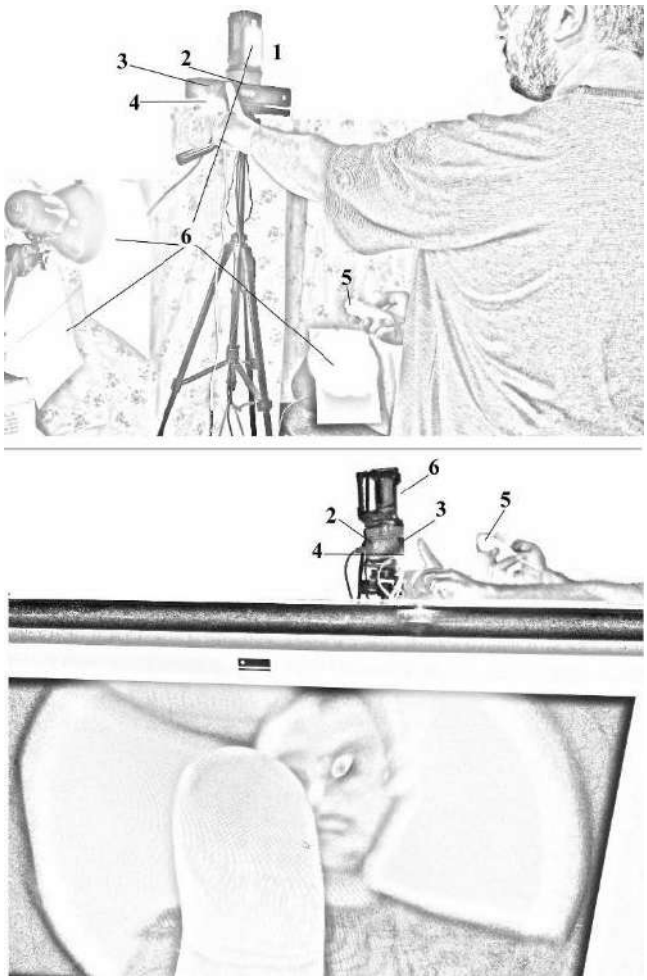


Figure 1. Experiment enviroment

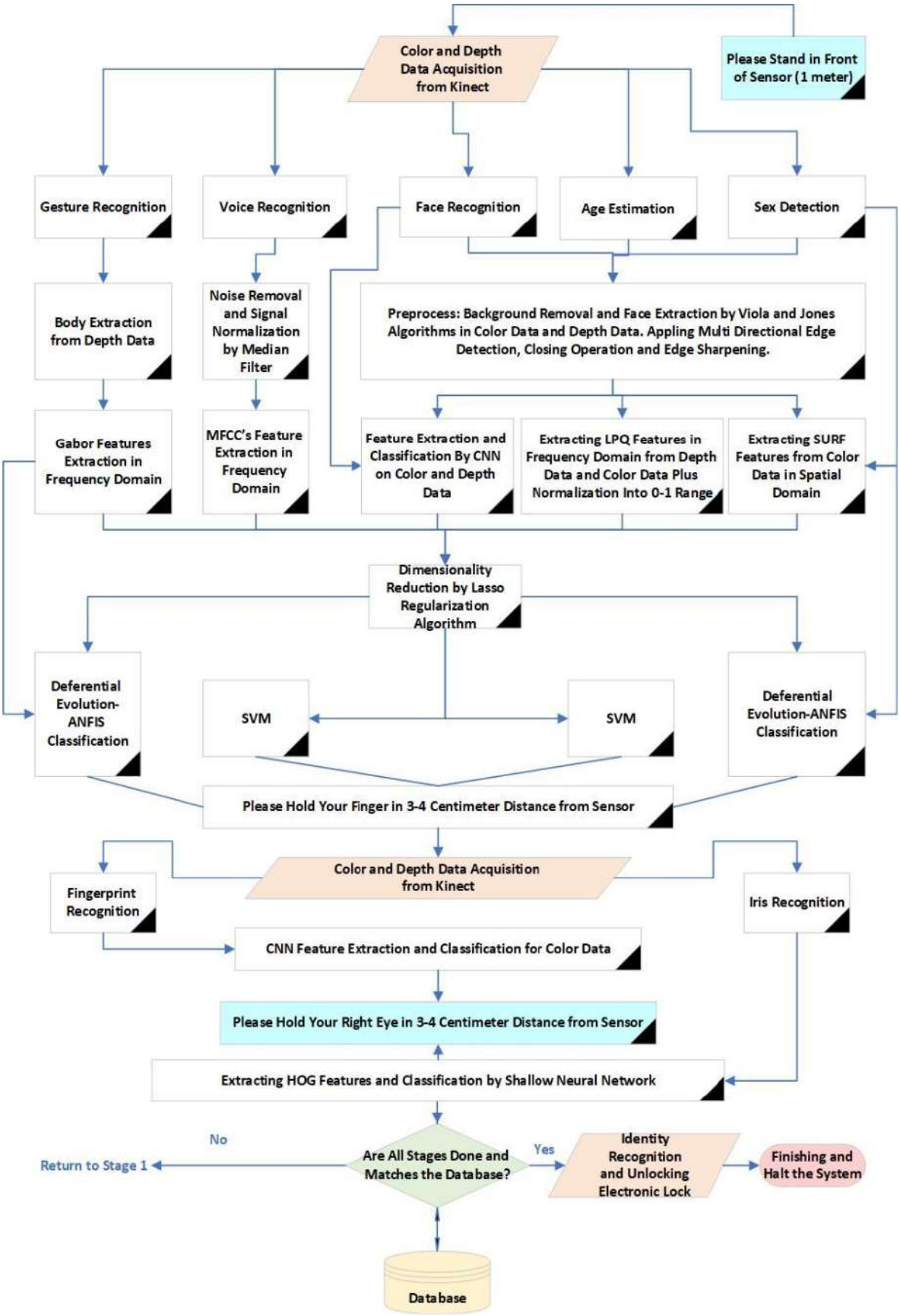


Figure 2. Proposed method flowchart

Figure 3 shows equipment used in the experiment. Also, Figuree 4 represents some steps of fingerprint and iris recognition steps. Figures 5 and 6 present different steps of the process (part 1 and 2).

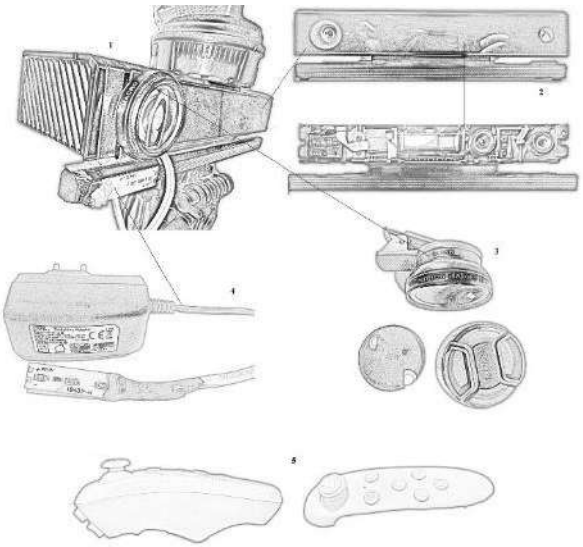


Figure 3. Experiment's Equepments

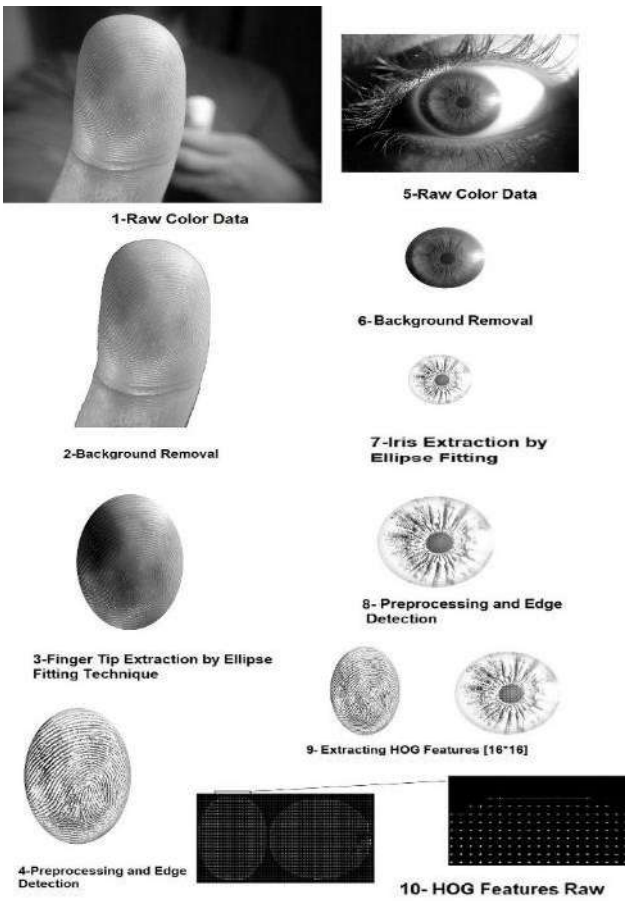


Figure 4. Some steps of fingerprint and iris recognition

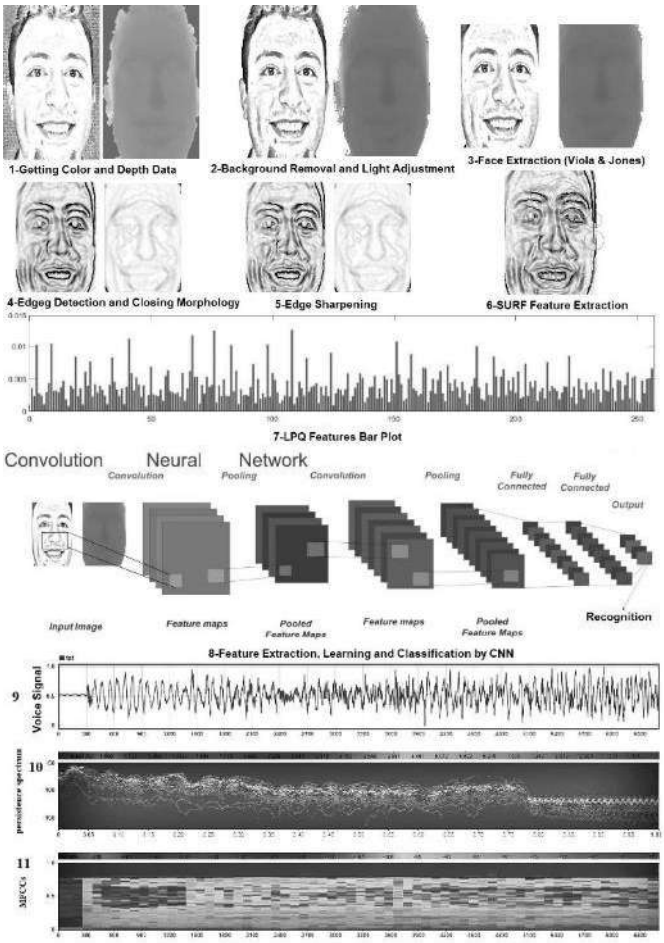


Figure 5. Different steps of the process (part 1)

IV. VALIDATION AND RESULTS

Experiment conducted on 6 subjects which their color and depth data of face, gesture, fingerprint, iris plus voice data is collected as a database for the system. Each subject has unique 5 second command for unlocking in voice part. Classification accuracy for CNN, SVM and DE-ANFIS for each subject and in whole is calculated and visible in Table I. Also, accuracy, precision and recall [23, 25] values are calculated for more details on the system's performance and is presented in the Table II. Figure 7 represents accuracy, precision and recall structure. Table III, IV and V are presenting confusion matrix for SVM, CNN and DE-ANFIS classification results, respectively

TABLE I. CLASSIFICATION RESULTS

	SVM	CNN	DE-ANFIS
Subject 1	98.12 %	99.60 %	98.03 %
Subject 2	98.99 %	99.09 %	97.52 %
Subject 3	98.14 %	99.73 %	97.47 %
Subject 4	97.98 %	99.16 %	97.63 %
Subject 5	99.11 %	99.98 %	98.19 %
Subject 6	98.68 %	98.90 %	98.63 %
Whole	**98.50 %**	**99.41 %**	**97.91 %**

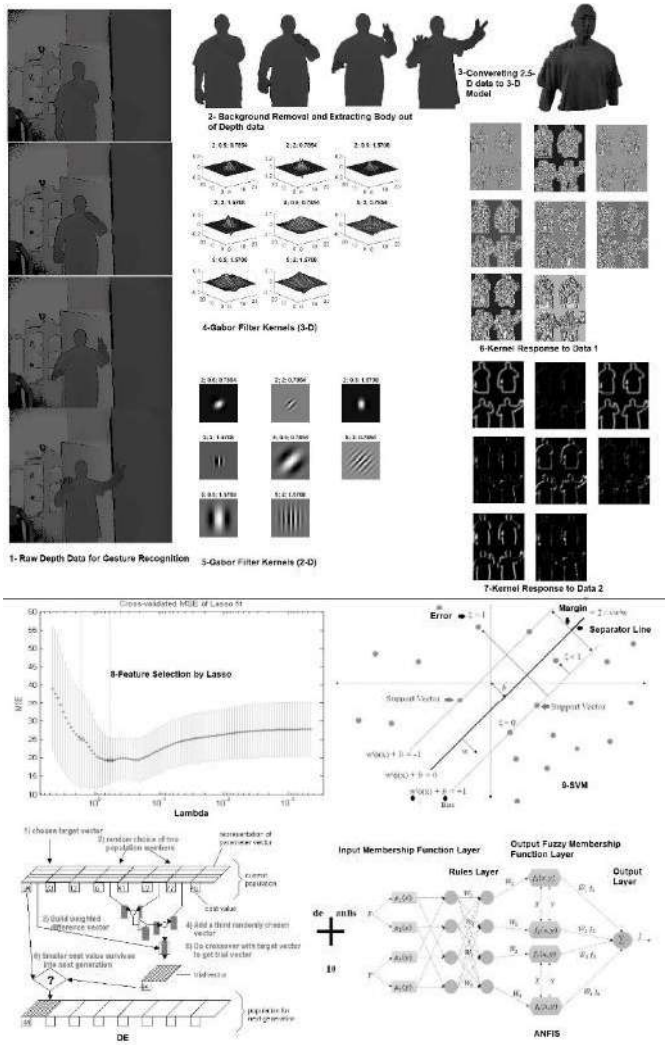


Figure 6. Different steps of the process (part 2)

$$\text{Precision} = \frac{\text{True Positive}}{\text{Actual Results}} \quad \text{or} \quad \frac{\text{True Positive}}{\text{True Positive + False Positive}}$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{Predicted Results}} \quad \text{or} \quad \frac{\text{True Positive}}{\text{True Positive + False Negative}}$$

$$\text{Accuracy} = \frac{\text{True Positive + True Negative}}{\text{Total}}$$

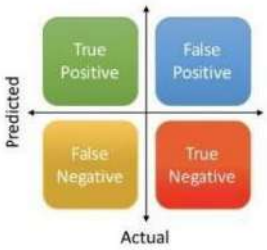


Figure 7. Accuracy, precision and recall

TABLE II. ACCURACY, PRECISION AND RECALL

Threshold	**Accuracy**	**Precision**	**Recall**
30	**99.12 %**	**99.19 %**	**98.73 %**
50	99.01 %	99.07 %	98.52 %
80	98.94 %	99.00 %	98.38 %

TABLE III. SVM CONFUSION MATRIX

	S 1	**S 2**	**S 3**	**S 4**	**S 5**	**S 6**
S 1	98 %	-	1	-	1	-
S 2	-	98 %	2	-	-	-
S 3	1	1	98%	-	-	-
S 4	-	-	1	97 %	2	-
S 5	1	-	-	-	99 %	-
S 6	-	-	-	-	2	98 %

TABLE IV. CNN CONFUSION MATRIX

	S 1	**S 2**	**S 3**	**S 4**	**S 5**	**S 6**
S 1	99 %	-	1	-	-	-
S 2	1	99 %	-	-	-	-
S 3	-	-	99 %	1	-	-
S 4	-	-	-	99 %	-	1
S 5	-	1	-	-	99 %	-
S 6	1	-	-	-	-	98 %

TABLE V. DE-ANFIS CONFUSION MATRIX

	S 1	**S 2**	**S 3**	**S 4**	**S 5**	**S 6**
S 1	98 %	-	-	-	-	2
S 2	-	97 %	1	1	-	1
S 3	1	1	97 %	1	-	-
S 4	-	-	-	97 %	3	-
S 5	-	-	-	-	98 %	2
S 6	-	-	2	-	-	98 %

As it is clear in Table I, all three classification algorithms achieved pretty acceptable recognition accuracy by the system. However, best result belongs to CNN, weakest to DE-ANFIS and SVM in the middle. Figure 7 shows accuracy, precision and recall's structure. Figure 8 presents the classification result for classifiers and Figure 9 represents accuracy, precision and recall results in graphical form.

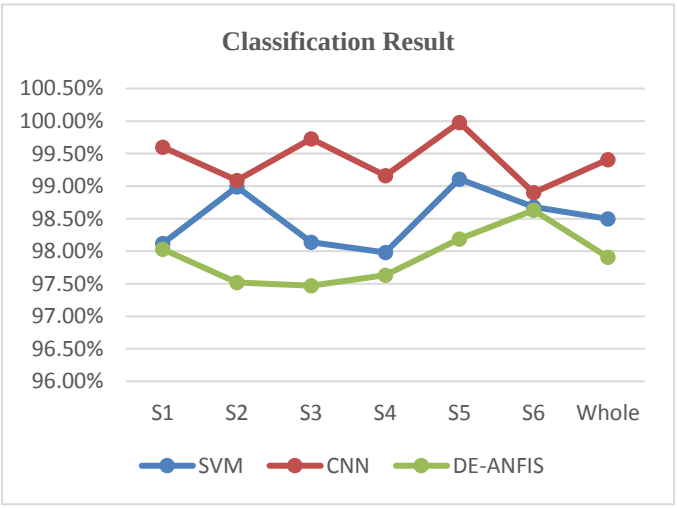


Figure 8. Classification Bar Plot

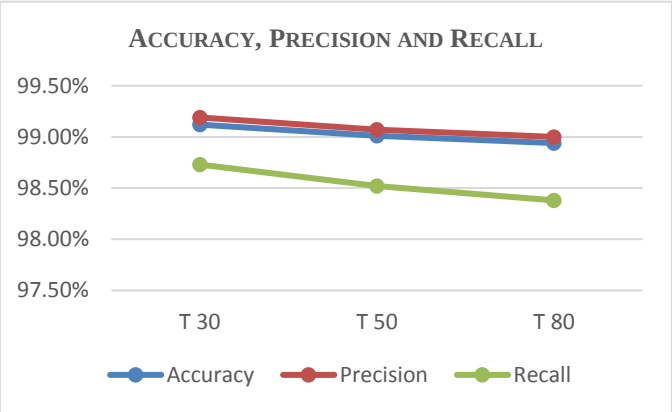


Figure 9. Accuracy, Precition and Recall Bar Plot

V. CONCLUSION AND SUGGESTIONS

Getting as much as possible potential that artificial intelligence offers in this expert system, made one of the most precise identity recognition systems for security systems as cheap as possible. Combining different algorithms in variety of tasks plus employing a suitable sensor brought the performance into a level higher. Overall performance of the system is very desirable and enhanced the era of learning-based security and surveillance systems. It is suggested to employ the system with other classification algorithms such as K-Nearest Neighborhood (K-NN), Tree and ensemble. Also, it is suggested to increase the number of subjects in the database. Extracting Scale-invariant feature transform (SIFT) and Local Binary Patterns (LBP) features and adding 3-D face recognition capability is of future works.

REFERENCES

[1] Mousavi, Seyed Muhammad Hossein, Vyacheslav Lyashenko, and Surya Prasath. "Analysis of a robust edge detection system in different color spaces using color and depth images." Компьютерная оптика 43.4 (2019).

[2] Mousavi, Seyed Muhammad Hossein, VB Surya Prasath, and Seyed Muhammad Hassan Mousavi. "Persian classical music instrument recognition (PCMIR) using a novel Persian music database." 2019 9th International Conference on Computer and Knowledge Engineering (ICCKE). IEEE, 2019.

[3] Dezfoulian, Mir Hossein, et al. "Optimization of the Ho-Kashyap classification algorithm using appropriate learning samples." 2016 Eighth International Conference on Information and Knowledge Technology (IKT). IEEE, 2016.

[4] Mousavi, Seyed Muhammad Hossein, S. Younes MiriNezhad, and Atiye Mirmoini. "A new support vector finder method, based on triangular calculations and K-means clustering." 2017 9th International Conference on Information and Knowledge Technology (IKT). IEEE, 2017.

[5] Mousavi, Seyed Muhammad Hossein, et al. "A PSO fuzzy-expert system: As an assistant for specifying the acceptance by NOET measures, at PH. D level." 2017 Artificial Intelligence and Signal Processing Conference (AISP). IEEE, 2017.

[6] Mousavi, Seyed Muhammad Hossein, Vincent Charles, and Tatiana Gherman. "An Evolutionary Pentagon Support Vector Finder Method." Expert Systems with Applications 150 (2020): 113284.

[7] Hu, Hao-Chun, et al. "Deep Learning Application for Vocal Fold Disease Prediction Through Voice Recognition: Preliminary Development Study." Journal of Medical Internet Research 23.6 (2021): e25247.

[8] Abozaid, Anter, et al. "Multimodal biometric scheme for human authentication technique based on voice and face recognition fusion." Multimedia Tools and Applications 78.12 (2019): 16345-16361.

[9] Peng, Chang, Mengyue Chen, and Xiaoning Jiang. "Under-display ultrasonic fingerprint recognition with finger vessel imaging." IEEE Sensors Journal 21.6 (2021): 7412-7419.

[10] Omran, Maryim, and Ebtesam N. AlShemmary. "An iris recognition system using deep convolutional neural network." Journal of Physics: Conference Series. Vol. 1530. No. 1. IOP Publishing, 2020.

[11] Liu, Fenglin, et al. "Kinect-based hand gesture recognition using trajectory information, hand motion dynamics and neural networks." Artificial Intelligence Review 52.1 (2019): 563-583.

[12] Mousavi, Seyed Muhammad Hossein. "A new way to age estimation for rgb-d images, based on a new face detection and extraction method for depth images." International Journal of Image, Graphics and Signal Processing 10 (2018): 10.

[13] https://patents.google.com/patent/US20060280346

[14] https://patents.google.com/patent/US20130089245A1/en

[15] Mousavi, Seyed Muhammad Hossain. "A New Clustering Method Using Evolutionary Algorithms for Determining Initial States, and Diverse Pairwise Distances for Clustering.", International Journal of Mechatronics, Electrical and Computer Technology (IJMEC), 9.31, (4098-4110), 2019.

[16] Hossein, Seyed Muhammad, et al. "Utilizing SURF Features and KLT Tracking Algorithm in Augmented Reality (AR), Using Kinect V. 2 with the Aim of Autism Therapy." Proceedings of the National Conference on Computer, Information Technology and Applications of Artificial Intelligence, Ahvaz, Iran. Vol. 5. 2020.

[17] Mousavi, Seyed Muhammad Hossein, and S. Younes Mirinezhad. "Iranian kinect face database (IKFDB): a color-depth based face database collected by kinect v. 2 sensor." SN Applied Sciences 3.1 (2021): 1-17.

[18] Gonzalez, Rafael C., and Paul Wintz. "Digital image processing(Book)." Reading, Mass., Addison-Wesley Publishing Co., Inc.(Applied Mathematics and Computation 13 (1977): 451.

[19] Viola, Paul, and Michael Jones. "Rapid object detection using a boosted cascade of simple features." Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001. Vol. 1. Ieee, 2001.

[20] Mousavi, S. Muhammad Hossein, and Marwa Kharazi. "An edge detection system for polluted images by gaussian, salt and pepper, poisson and speckle noises." 4th National Conference on Information Technology, Computer & TeleCommunication. 2017.

[21] MOUSAVI, SEYED MUHAMMAD HOSSEIN. "A SPATIAL AND FREQUENCY BASED METHOD FOR MICRO FACIAL EXPRESSIONS RECOGNITION USING COLOR AND DEPTH IMAGES. Journal of Software Engineering & Intelligent Systems 6 (1), 17, 2021".

[22] Mousavi S. M. H., MiriNezhad S. Y., Lyashenko V. An Evolutionary-Based Adaptive Neuro-Fuzzy Expert System as a Family Counselor before Marriage with the Aim of Divorce Rate Reduction // Conference: 2nd International Conference on Research Knowledge Base in Computer Engineering and IT, At Tehran-Iran. – 2017. – P. 1-10.

[23] Mousavi, Seyed Muhammad Hossein, and S. Younes Mirinezhad. "Fingerprint Recognition Out of Object's Surface Using Ordinary Cameras with the Aid of Magnifier." Recall (without magnifier), 3rd National Conference on Computer, information technology and applications of artificial intelligence, ahvaz, iran, (2020), 30.89.6: 88-3.

[24] Razavi-Termeh, Seyed Vahid, Kourosh Shirani, and Mehrdad Pasandi. "Mapping of landslide susceptibility using the combination of neuro-fuzzy inference system (ANFIS), ant colony (ANFIS-ACOR), and differential evolution (ANFIS-DE) models." Bulletin of Engineering Geology and the Environment 80.3 (2021): 2045-2067.

[25] Polyvyanyy, Artem, et al. "Monotone precision and recall measures for comparing executions and specifications of dynamic systems." ACM Transactions on Software Engineering and Methodology (TOSEM) 29.3 (2020): 1-41.