RECENT ADVANCES in NEURAL NETWORKS and FUZZY SYSTEMS

Proceedings of the 2014 International Conference on Neural Networks - Fuzzy Systems (NN-FS '14)

Venice, Italy
March 15-17, 2014

RECENT ADVANCES in NEURAL NETWORKS and FUZZY SYSTEMS

Proceedings of the 2014 International Conference on Neural Networks - Fuzzy Systems (NN-FS '14)

Venice, Italy
March 15-17, 2014

All papers of the present volume were peer reviewed by no less than two independent reviewers. Acceptance was granted when both reviewers' recommendations were positive.

Series: Recent Advances in Computer Engineering Series - 19

ISSN: 1790-5109
ISBN: 978-1-61804-227-9

RECENT ADVANCES in NEURAL NETWORKS and FUZZY SYSTEMS

Proceedings of the 2014 International Conference on Neural Networks - Fuzzy Systems (NN-FS '14)

Venice, Italy
March 15-17, 2014

Organizing Committee

General Chairs (EDITORS)

- Prof. Yingxu Wang, PhD,
  P.Eng, F.WIF, F.ICIC, SM.IEEE, SM.ACM
  Schulich School of Engineering
  University of Calgary
  Calgary, AB, Canada T2N 1N4
- Prof. Pierre Borne, IEEE France Section Chair,
  IEEE Fellow, IEEE/SMC Past President, Ecole Centrale de Lille
  BP 48, 59651 Villeneuve d'Ascq, France
  Prof. Imre Rudas, Obuda University, Budapest, Hungary

Senior Program Chair

- Prof. Janusz Kacprzyk,
  IEEE Fellow,
  Polish Academy of Sciences, Warsaw, Poland

Program Chairs

- Prof. Filippo Neri
  Dipartimento di Informatica e Sistemistica
  University of Naples "Federico II"
  Naples, Italy
- Prof. Constantin Udriste,
  University Politehnica of Bucharest,
  Bucharest, Romania
- Prof. Sandra Sendra
  Instituto de Inv. para la Gestión Integrada de Zonas Costeras (IGIC)
  Universidad Politécnica de Valencia
  Spain

Tutorials Chair

- Professor Pradip Majumdar
  Department of Mechanical Engineering
  Northern Illinois University
  Dekalb, Illinois, USA

Special Session Chair

- Prof. Pavel Varacha
  Tomas Bata University in Zlin
  Faculty of Applied Informatics
  Department of Informatics and Artificial Intelligence
  Zlin, Czech Republic

Workshops Chair

- Prof. Ryszard S. Choras
  Institute of Telecommunications
  University of Technology & Life Sciences
  Bydgoszcz, Poland

Local Organizing Chair

- Assistant Prof. Klimis Ntalianis,
  Tech. Educ. Inst. of Athens (TEI),
  Athens, Greece

Publication Chair

- Prof. Gen Qi Xu
  Department of Mathematics
  Tianjin University
  Tianjin, China

Publicity Committee

- Prof. Reinhard Neck
  Department of Economics
  Klagenfurt University
  Klagenfurt, Austria
- Prof. Myriam Lazard
  Institut Superieur d' Ingenierie de la Conception
  Saint Die, France

International Liaisons

- Prof. Ka-Lok Ng
  Department of Bioinformatics
  Asia University
  Taichung, Taiwan
- Prof. Olga Martin
  Applied Sciences Faculty
  Politehnica University of Bucharest
  Romania
- Prof. Vincenzo Niola
  Departement of Mechanical Engineering for Energetics
  University of Naples "Federico II"
  Naples, Italy
- Prof. Eduardo Mario Dias
  Electrical Energy and Automation
  Engineering Department
  Escola Politecnica da Universidade de Sao Paulo
  Brazil

Steering Committee

- Professor Aida Bulucea, University of Craiova, Romania
- Professor Zoran Bojkovic, Univ. of Belgrade, Serbia
- Prof. Metin Demiralp, Istanbul Technical University, Turkey
- Professor Imre Rudas, Obuda University, Budapest, Hungary

Program Committee for NEURAL NETWORKS

Prof. Pierre Borne, IEEE France Section Chair, IEEE Fellow, IEEE/SMC Past President, Ecole Centrale de Lille BP 48, 59651 Villeneuve d'Ascq, France
Prof. Leon Chua, IEEE Fellow,University of Berkeley, USA
Prof. Narsingh Deo, IEEE Fellow, Department of Computer Science, University of Central Florida, Orlando, FL, USA
Prof. Wasfy B Mikhael, IEEE Fellow, University of Central Florida Orlando, USA
Prof. Panos Pardalos, IEEE Fellow, University of Florida, USA
Prof. Jim Austin, Computer Science, University of York, York, UK
Prof. Stefanos Kollias, National Technical University of Athens, Athens, Greece
Prof. Bruno Apolloni (President of the Italian Society of Neural Networks - SIREN), University of Milan, Milan, Italy
Prof. Alessandro Di Nuovo, Plymouth University, Plymouth, UK
Prof. Thomas Wennekers, Plymouth University, Plymouth, UK
Prof. Dominic Palmer-Brown, London Metropolitan University, London, UK
Prof. Yi Ming Zou, University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA
Prof. Sebastian Pannasch, Technische Universitaet Dresden, Dresden, Germany
Prof. Yutaka Maeda, Faculty of Engineering Science, Kansai University, Japan
Prof. Andreas Koenig, Technische Universitaet Kaiserslautern, D-67663 Kaiserslautern, Germany
Prof. Jiann-Ming Wu, National Dong Hwa University, Hualien, Taiwan
Prof. Paolo Gastaldo, University of Genova, Genova, Italy
Prof. Friedhelm Schwenker, Universitat Ulm, 89069 Ulm, Germany
Prof. Juan Ignacio Arribas, Universidad de Valladolid, Valladolid, Spain
Prof. Tienfuan Kerh, National Pingtung University of Science and Technology, Pingtung, Taiwan
Prof. Florin Gorunescu, University of Medicine and Pharmacy of Craiova, Craiova, Romania
Prof. Imre J. Rudas, Obuda University, Budapest, Hungary
Prof. Francesco Camastra, University of Naples - Parthenope, Naples, Italy
Prof. Kyong Joo Oh, Yonsei University, Korea
Prof. Francesco Marcelloni, University of Pisa, Pisa, Italy
Prof. Claudio Mirasso, Universitat de les Illes Balears, Palma de Mallorca, Spain
Prof. Günther Palm, Ulm University, Ulm, Germany
Prof. Giorgio Valentini, University of Milan (Milano), Italy
Prof. Zhiyuan Luo, Royal Holloway, University of London, Egham, Surrey, UK
Prof. Alessio Micheli, Universita di Pisa, Pisa, Italy
Prof. Ping-Feng Pai, National Chi Nan University, Puli, Nantou, Taiwan
Prof. Manwai Mak, Hong Kong Polytechnic University, Hung Hom, Hong Kong
Prof. Alexander Gegov, University of Portsmouth, Portsmouth, UK
Prof. Anastasios Tefas, Aristotle University of Thessaloniki, Thessaloniki, Greece
Prof. Hazem M. El-Bakry, Mansoura University, Mansoura, Egypt
Prof. Alessandro Rozza, University of Naples - Parthenope, Naples, Italy
Prof. Ryszard S. Choras, University of Technology & Life Sciences, Bydgoszcz, Poland
Prof. Dietrich Klakow, Saarland University, Saarbrucken, Germany
Prof. George Mengov, Sofia University St. Kliment Ohridski, Sofia, Bulgaria
Prof. Mirko Novak, Czech Technical University in Prague, Czech Republic
Dr. Tonatiuh Pena Centeno, Universitat of Greifswald, Greifswald, Germany
Dr. Mauro Gaggero, Institute of Intelligent Systems for Automation (ISSIA), National Research Council of Italy, Genoa, Italy

Program Committee for FUZZY SYSTEMS

Prof. Lotfi Zadeh, IEEE Fellow, University of Berkeley, USA
Prof. Michio Sugeno, IEEE Fellow, Tokyo Institute of Technology, Japan
Prof. Janusz Kacprzyk, IEEE Fellow, Polish Academy of Sciences, Warsaw, Poland
Prof. Pierre Borne, IEEE France Section Chair, IEEE Fellow, IEEE/SMC Past President, Ecole Centrale de Lille, France

Additional Reviewers

Angel F. Tenorio	Universidad Pablo de Olavide, Spain
Ole Christian Boe	Norwegian Military Academy, Norway
Abelha Antonio	Universidade do Minho, Portugal
Xiang Bai	Huazhong University of Science and Technology, China
Genqi Xu	Tianjin University, China
Moran Wang	Tsinghua University, China
Minhui Yan	Shanghai Maritime University, China
Jon Burley	Michigan State University, MI, USA
Shinji Osada	Gifu University School of Medicine, Japan
Bazil Taha Ahmed	Universidad Autonoma de Madrid, Spain
Konstantin Volkov	Kingston University London, UK
Tetsuya Shimamura	Saitama University, Japan
George Barreto	Pontificia Universidad Javeriana, Colombia
Tetsuya Yoshida	Hokkaido University, Japan
Deolinda Rasteiro	Coimbra Institute of Engineering, Portugal
Matthias Buyle	Artesis Hogeschool Antwerpen, Belgium
Dmitrijs Serdjuks	Riga Technical University, Latvia
Kei Eguchi	Fukuoka Institute of Technology, Japan
Imre Rudas	Obuda University, Budapest, Hungary
Francesco Rotondo	Polytechnic of Bari University, Italy
Valeri Mladenov	Technical University of Sofia, Bulgaria
Andrey Dmitriev	Russian Academy of Sciences, Russia
James Vance	The University of Virginia's College at Wise, VA, USA
Masaji Tanaka	Okayama University of Science, Japan
Sorinel Oprisan	College of Charleston, CA, USA
Hessam Ghasemnejad	Kingston University London, UK
Santoso Wibowo	CQ University, Australia
M. Javed Khan	Tuskegee University, AL, USA
Manoj K. Jha	Morgan State University in Baltimore, USA
Miguel Carriegos	Universidad de Leon, Spain
Philippe Dondon	Institut polytechnique de Bordeaux, France
Kazuhiko Natori	Toho University, Japan
Jose Flores	The University of South Dakota, SD, USA
Takuya Yamano	Kanagawa University, Japan
Frederic Kuznik	National Institute of Applied Sciences, Lyon, France
Lesley Farmer	California State University Long Beach, CA, USA
João Bastos	Instituto Superior de Engenharia do Porto, Portugal
Zhong-Jie Han	Tianjin University, China
Francesco Zirilli	Sapienza Universita di Roma, Italy
Yamagishi Hiromitsu	Ehime University, Japan
Eleazar Jimenez Serrano	Kyushu University, Japan
Alejandro Fuentes-Penna	Universidad Autónoma del Estado de Hidalgo, Mexico
José Carlos Metrôlho	Instituto Politecnico de Castelo Branco, Portugal
Stavros Ponis	National Technical University of Athens, Greece

Table of Contents

ISBN: 978-1-61804-227-9

Keynote Lecture 1

On the Distinguished Role of the Mittag-Leffler and Wright Functions in Fractional Calculus

Professor Francesco Mainardi
Department of Physics, University of Bologna, and INFN
Via Irnerio 46, I-40126 Bologna, Italy
E-mail: francesco.mainardi@bo.infn.it.it

Abstract: Fractional calculus, in allowing integrals and derivatives of any positive real order (the term "fractional" is kept only for historical reasons), can be considered a branch of mathematical analysis which deals with integro-di erential equations where the integrals are of convolution type and exhibit (weakly singular) kernels of power-law type. As a matter of fact fractional calculus can be considered a laboratory for special functions and integral transforms. Indeed many problems dealt with fractional calculus can be solved by using Laplace and Fourier transforms and lead to analytical solutions expressed in terms of transcendental functions of Mittag-Leffler and Wright type. In this plenary lecture we discuss some interesting problems in order to single out the role of these functions. The problems include anomalous relaxation and diffusion and also intermediate phenomena.

Brief Biography of the Speaker: For a full biography, list of references on author's papers and books see:
Home Page: http://www.fracalmo.org/mainardi/index.htm
and http://scholar.google.com/citations?user=UYxWyEEAAAAJ&hl=en&oi=ao

ISBN: 978-1-61804-227-9

Keynote Lecture 2

Latest Advances in Neuroinformatics and Fuzzy Systems

Yingxu Wang, PhD, Prof., PEng, FWIF, FICIC, SMIEEE, SMACM
President, International Institute of Cognitive Informatics and Cognitive Computing (ICIC)
Director, Laboratory for Cognitive Informatics and Cognitive Computing
Dept. of Electrical and Computer Engineering
Schulich School of Engineering
University of Calgary
2500 University Drive NW,
Calgary, Alberta, Canada T2N 1N4
E-mail: yingxu@ucalgary.ca

Abstract: Investigations into the neurophysiological foundations of neural networks in neuroinformatics [Wang, 2013] have led to a set of rigorous mathematical models of neurons and neural networks in the brain using contemporary denotational mathematics [Wang, 2008, 2012]. A theory of neuroinformatics is recently developed for explaining the roles of neurons in internal information representation, transmission, and manipulation [Wang & Fariello, 2012]. The formal neural models reveal the differences of structures and functions of the association, sensory and motor neurons. The pulse frequency modulation (PFM) theory of neural networks [Wang & Fariello, 2012] is established for rigorously analyzing the neurosignal systems in complex neural networks. It is noteworthy that the Hopfield model of artificial neural networks [Hopfield, 1982] is merely a prototype closer to the sensory neurons, though the majority of human neurons are association neurons that function significantly different as the sensory neurons. It is found that neural networks can be formally modeled and manipulated by the neural circuit theory [Wang, 2013]. Based on it, the basic structures of neural networks such as the serial, convergence, divergence, parallel, feedback circuits can be rigorously analyzed. Complex neural clusters for memory and internal knowledge representation can be deduced by compositions of the basic structures.

Fuzzy inferences and fuzzy semantics for human and machine reasoning in fuzzy systems [Zadeh, 1965, 2008], cognitive computers [Wang, 2009, 2012], and cognitive robots [Wang, 2010] are a frontier of cognitive informatics and computational intelligence. Fuzzy inference is rigorously modeled in inference algebra [Wang, 2011], which recognizes that humans and fuzzy cognitive systems are not reasoning on the basis of probability of causations rather than formal algebraic rules. Therefore, a set of fundamental fuzzy operators, such as those of fuzzy causality as well as fuzzy deductive, inductive, abductive, and analogy rules, is formally elicited. Fuzzy semantics is quantitatively modeled in semantic algebra [Wang, 2013], which formalizes the qualitative semantics of natural languages in the categories of nouns, verbs, and modifiers (adjectives and adverbs). Fuzzy semantics formalizes nouns by concept algebra [Wang, 2010],

verbs by behavioral process algebra [Wang, 2002, 2007], and modifiers by fuzzy semantic algebra [Wang, 2013]. A wide range of applications of fuzzy inference, fuzzy semantics, neuroinformatics, and denotational mathematics have been implemented in cognitive computing, computational intelligence, fuzzy systems, cognitive robotics, neural networks, neurocomputing, cognitive learning systems, and artificial intelligence.

Brief Biography of the Speaker: Yingxu Wang is professor of cognitive informatics and denotational mathematics, President of International Institute of Cognitive Informatics and Cognitive Computing (ICIC, http://www.ucalgary.ca/icic/) at the University of Calgary. He is a Fellow of ICIC, a Fellow of WIF (UK), a P.Eng of Canada, and a Senior Member of IEEE and ACM. He received a PhD in software engineering from the Nottingham Trent University, UK, and a BSc in Electrical Engineering from Shanghai Tiedao University. He was a visiting professor on sabbatical leaves at Oxford University (1995), Stanford University (2008), University of California, Berkeley (2008), and MIT (2012), respectively. He is the founder and steering committee chair of the annual IEEE International Conference on Cognitive Informatics and Cognitive Computing (ICCI*CC) since 2002. He is founding Editor-in-Chief of International Journal of Cognitive Informatics and Natural Intelligence (IJCINI), founding Editor-in-Chief of International Journal of Software Science and Computational Intelligence (IJSSCI), Associate Editor of IEEE Trans. on SMC (Systems), and Editor-in-Chief of Journal of Advanced Mathematics and Applications (JAMA). Dr. Wang is the initiator of a few cutting-edge research fields or subject areas such as denotational mathematics, cognitive informatics, abstract intelligence (□I), cognitive computing, software science, and basic studies in cognitive linguistics. He has published over 160 peer reviewed journal papers, 230+ peer reviewed conference papers, and 25 books in denotational mathematics, cognitive informatics, cognitive computing, software science, and computational intelligence. He is the recipient of dozens international awards on academic leadership, outstanding contributions, best papers, and teaching in the last three decades.
http://www.ucalgary.ca/icic/
http://scholar.google.ca/citations?user=gRVQjskAAAAJ&hl=en

Editor-in-Chief, International Journal of Cognitive Informatics and Natural Intelligence
Editor-in-Chief, International Journal of Software Science and Computational Intelligence
Associate Editor, IEEE Transactions on System, Man, and Cybernetics - Systems
Editor-in-Chief, Journal of Advanced Mathematics and Applications
Chair, The Steering Committee of IEEE ICCI*CC Conference Series

Keynote Lecture 3

Recent Advances and Future Trends on Atomic Engineering of III-V Semiconductor for Quantum Devices from Deep UV (200nm) up to THZ (300 microns)

Professor Manijeh Razeghi
Center for Quantum Devices
Department of Electrical Engineering and Computer Science
Northwestern University
Evanston, Illinois 60208
USA
E-mail: razeghi@eecs.northwestern.edu

Abstract: Nature offers us different kinds of atoms, but it takes human intelligence to put them together in an elegant way in order to realize functional structures not found in nature. The so-called III-V semiconductors are made of atoms from columns III (B, Al, Ga, In. Tl) and columns V(N, As, P, Sb,Bi) of the periodic table, and constitute a particularly rich variety of compounds with many useful optical and electronic properties. Guided by highly accurate simulations of the electronic structure, modern semiconductor optoelectronic devices are literally made atom by atom using advanced growth technology such as Molecular Beam Epitaxy (MBE) and Metal Organic Chemical Vapor Deposition (MOCVD). Recent breakthroughs have brought quantum engineering to an unprecedented level, creating light detectors and emitters over an extremely wide spectral range from 0.2 mm to 300 mm. Nitrogen serves as the best column V element for the short wavelength side of the electromagnetic spectrum, where we have demonstrated III-nitride light emitting diodes and photo detectors in the deep ultraviolet to visible wavelengths. In the infrared, III-V compounds using phosphorus ,arsenic and antimony from column V ,and indium, gallium, aluminum, ,and thallium from column III elements can create interband and intrsuband lasers and detectors based on quantum-dot (QD) or type-II superlattice (T2SL). These are fast becoming the choice of technology in crucial applications such as environmental monitoring and space exploration. Last but not the least, on the far-infrared end of the electromagnetic spectrum, also known as the terahertz (THz) region, III-V semiconductors offer a unique solution of generating THz waves in a compact device at room temperature. Continued effort is being devoted to all of the above mentioned areas with the intention to develop smart technologies that meet the current challenges in environment, health, security, and energy. This talk will highlight my contributions to the world of III-V semiconductor Nano scale optoelectronics. Devices from deep UV-to THz.

Brief Biography of the Speaker: Manijeh Razeghi received the Doctorat d'État es Sciences Physiques from the Université de Paris, France, in 1980.
After heading the Exploratory Materials Lab at Thomson-CSF (France), she joined Northwestern University, Evanston, IL, as a Walter P. Murphy Professor and Director of the Center for

ISBN: 978-1-61804-227-9

Quantum Devices in Fall 1991, where she created the undergraduate and graduate program in solid-state engineering. She is one of the leading scientists in the field of semiconductor science and technology, pioneering in the development and implementation of major modern epitaxial techniques such as MOCVD, VPE, gas MBE, and MOMBE for the growth of entire compositional ranges of III-V compound semiconductors. She is on the editorial board of many journals such as Journal of Nanotechnology, and Journal of Nanoscience and Nanotechnology, an Associate Editor of Opto-Electronics Review. She is on the International Advisory Board for the Polish Committee of Science, and is an Adjunct Professor at the College of Optical Sciences of the University of Arizona, Tucson, AZ. She has authored or co-authored more than 1000 papers, more than 30 book chapters, and fifteen books, including the textbooks Technology of Quantum Devices (Springer Science+Business Media, Inc., New York, NY U.S.A. 2010) and Fundamentals of Solid State Engineering, 3rd Edition (Springer Science+Business Media, Inc., New York, NY U.S.A. 2009). Two of her books, MOCVD Challenge Vol. 1 (IOP Publishing Ltd., Bristol, U.K., 1989) and MOCVD Challenge Vol. 2 (IOP Publishing Ltd., Bristol, U.K., 1995), discuss some of her pioneering work in InP-GaInAsP and GaAs-GaInAsP based systems. The MOCVD Challenge, 2nd Edition (Taylor & Francis/CRC Press, 2010) represents the combined updated version of Volumes 1 and 2. She holds 50 U.S. patents and has given more than 1000 invited and plenary talks. Her current research interest is in nanoscale optoelectronic quantum devices.

Dr. Razeghi is a Fellow of MRS, IOP, IEEE, APS, SPIE, OSA, Fellow and Life Member of Society of Women Engineers (SWE), Fellow of the International Engineering Consortium (IEC), and a member of the Electrochemical Society, ACS, AAAS, and the French Academy of Sciences and Technology. She received the IBM Europe Science and Technology Prize in 1987, the Achievement Award from the SWE in 1995, the R.F. Bunshah Award in 2004, and many best paper awards.

ISBN: 978-1-61804-227-9

Fuzzy Semantic Models of Fuzzy Concepts in Fuzzy Systems

Yingxu Wang

***Abstract*—**The fuzzy properties of language semantics are a central problem towards machine-enabled natural language processing in cognitive linguistics, fuzzy systems, and computational linguistics. A formal method for rigorously describing and manipulating fuzzy semantics is sought for bridging the gap between humans and cognitive fuzzy systems. The mathematical model of fuzzy concepts is rigorously described as a hyperstructure of fuzzy sets of attributes, objects, relations, and qualifications, which serves as the basic unit of fuzzy semantics for denoting languages entities in semantic analyses. The formal fuzzy concept is extended to complex structures where fuzzy modifiers and qualifiers are considered. An algebraic approach is developed to manipulate composite fuzzy semantic as a deductive process from a fuzzy concept to the determined semantics. The denotational mathematical structure of fuzzy semantic inference not only explains the fuzzy nature of human semantics and its comprehension, but also enables cognitive machines and fuzzy systems to mimic the human fuzzy inference mechanisms in cognitive linguistics, cognitive computing, and computational intelligence.

***Keywords*—**Fuzzy systems, fuzzy semantics, fuzzy concept, cognitive linguistics, fuzzy inference, mathematical models of semantics, cognitive informatics, cognitive computing, soft computing, abstract intelligence, computational intelligence.

I. Introduction

FUZZY *semantics comprehension* and *fuzzy inference* are two of the central abilities of human brains that play a crucial role in thinking, perception, and problem solving [1, 2, 12, 13, 15, 17, 18, 23, 24, 25, 26, 28, 29, 31, 32, 35, 36]. *Semantics* in linguistics represents the meaning or the intension and extension of a language entity [3, 4, 7, 9, 10, 14]. *Formal semantics* [8, 9, 11, 19, 22, 27] focus on mathematical models for denoting meanings of symbols, concepts, functions, and behaviors, as well as their relations, which can be deduced onto a set of known concepts and behavioral processes in cognitive linguistics [5, 6, 28]. An *inference* is a cognitive process that deduces a proposition, particularly a causation, based on logical relations.

This work was supported in part by a discovery fund granted by the Natural Sciences and Engineering Research Council of Canada (NSERC).

Y. Wang is with the International Institute of Cognitive Informatics and Cognitive Computing (ICIC), Dept. of Electrical and Computer Engineering, Schulich School of Engineering, University of Calgary, 2500 University Drive NW, Calgary, Alberta, Canada T2N 1N4 (Tel: +1 403 220-6141, Fax: +1 403 282-6855, Email: yingxu@ucalgary.ca).

The taxonomy of semantics in natural languages can be classified into three categories [3, 4, 9, 14, 28, 32] known as those of entities (noun and noun phrases), behaviors (verbs and verb phrases), and modifiers (adjectives, adverbs, and related phrases). Semantics can also be classified into the categories of *to-be, to-have,* and *to-do* semantics [27]. A *to-be* semantics infers the meaning of an equivalent relation between an unknown and a known entity or concept. A *to-have* semantics denotes the meaning of a possessive structure or a composite entity. A *to-do* semantics embodies the process of a behavior or an action in a 5-dimensional behavioral space [22, 27].

The fuzzy nature of linguistic semantics as well as its cognition stems from inherent semantic ambiguity, context variability, and individual perceptions influenced by heterogeneous knowledge bases. Almost all problems in natural language processing and semantic analyses are constrained by these fundamental issues. Lotfi A. Zadeh extended methods for inferences to fuzzy sets and fuzzy logic [30, 33, 37], which provide a mathematical means for dealing with uncertainty and imprecision in reasoning, qualification, and quantification, particularly where vague linguistic variables are involved. Fuzzy inferences based on fuzzy sets are navel denotational mathematical means for rigorously dealing with degrees of matters, uncertainties, and vague semantics of linguistic entities, as well as for precisely reasoning the semantics of fuzzy causations. Typical fuzzy inference rules are those of fuzzy argument, implementation, deduction, induction, abduction, and analogy [22, 26, 32, 37].

This paper presents a theory of fuzzy concepts and fuzzy semantics for formal semantic manipulation in fuzzy systems and cognitive linguistics. The mathematical model of abstract fuzzy concepts is introduced in Section 2, which serves as the basic unit of fuzzy semantics in natural languages. A fuzzy concept is modeled as a fuzzy hyperstructure encompassing the fuzzy sets of attributes, objects, relations, and qualifications. Based on the mathematical model of fuzzy concepts, fuzzy semantic comprehension is reduced to a deduction process by algebraic operations on the fuzzy semantics. The mathematical model of fuzzy concept is extended to complex ones in Section 3 where fuzzy qualifiers are involved to modify fuzzy concepts. Algebraic operations on composite fuzzy semantics deduce the fuzzy semantics of a composite fuzzy concept to determined implications. The denotational mathematical structures of fuzzy semantic inference are elaborated by real-world examples in order to demonstrate their applications in cognitive linguistics, fuzzy systems, cognitive computing, and computational intelligence [6, 8, 12, 16, 20, 21].

II. FUZZY SEMANTICS OF CONCEPTS IN FUZZY INFERENCES

The semantics of an entity in a natural language is used to be vaguely represented by a noun or noun phrase. In order to rigorously express the intension and extension of an entity expressed by a word, the noun entities can be formally described by an abstract concept in *concept algebra* [19] and *semantic algebra* [27]. An abstract concept is a cognitive unit to identify and model a concrete entity in the physical world or an abstract object in the perceived world, which can be formally described as follows.

Definition 1. Let $\mathfrak{O}$ denote a finite nonempty set of *objects*, and $\mathfrak{A}$ be a finite nonempty set of *attributes*. The *semantic discourse* $\mathfrak{U}_s$ is a triple, i.e.:

$$\begin{aligned}\mathfrak{U}_s &= (\mathfrak{O}, \mathfrak{A}, \mathfrak{R}) \\ &= \mathfrak{R}: \mathfrak{O} \to \mathfrak{O} \,|\, \mathfrak{O} \to \mathfrak{A} \,|\, \mathfrak{A} \to \mathfrak{O} \,|\, \mathfrak{A} \to \mathfrak{A}\end{aligned} \tag{1}$$

where $\mathfrak{R}$ is a set of relations between $\mathfrak{O}$ and $\mathfrak{A}$.

On the basis of the *semantic discourse*, a formal fuzzy concept can be defined as a certain composition of subsets of the three kinds of elements known as the objects, attributes, and relations.

Definition 2. A *fuzzy concept* $\widetilde{C}$ is a hyperstructure of language entities denoted by a 5-tuple encompassing the fuzzy sets of attributes $\widetilde{A}$, objects $\widetilde{O}$, internal relations $\widetilde{R^i}$, external relations $\widetilde{R^o}$, and qualifications $\widetilde{Q}$ i.e.:

$$\widetilde{C} \triangleq (\widetilde{A}, \widetilde{O}, \widetilde{R^i}, \widetilde{R^o}, \widetilde{Q}) \tag{2}$$

where

- $\widetilde{A}$ is a fuzzy set of *attributes* as the *intension* of the concept $\widetilde{C}$:

$$\widetilde{A} = \{(a_1, \mu(a_1)), (a_2, \mu(a_2)), \ldots, (a_n, \mu(a_n))\} \subseteq Þ\mathfrak{A} \tag{3}$$

where $Þ\mathfrak{A}$ denotes a power set of $\mathfrak{A}$.

- $\widetilde{O}$ is a fuzzy set of *objects* as the *extension* of the concept $\widetilde{C}$:

$$\widetilde{O} = \{(o_1, \mu(o_1)), (o_2, \mu(o_2)), \ldots, (o_m, \mu(o_m))\} \subseteq Þ\mathfrak{O} \tag{4}$$

- $\widetilde{R^i}$ is a fuzzy set of *internal relations* between the fuzzy sets of objects $\widetilde{O}$ and attributes $\widetilde{A}$:

$$\begin{aligned}\widetilde{R^i} &= \widetilde{O} \times \widetilde{A} \subseteq Þ\mathfrak{R} \\ &= \mathop{R}_{j=1}^{|\widetilde{O}|} \mathop{R}_{i=1}^{|\widetilde{A}|} ((o_j, a_i), \mu(o_i) \bullet \mu(a_j))\end{aligned} \tag{5}$$

where the *big-R notation* [18, 25] expresses the Cartesian product of a series of repeated cross operations between o_j and a_i, $1 \leq j \leq m$ and $1 \leq i \leq n$, which results in all the (o_j, a_i) pairs.

- $\widetilde{R^o}$ is a fuzzy set of *external relations* between the fuzzy concept $\widetilde{C}$ and all potential ones $\widetilde{C'}$ in a knowledge base in $\mathfrak{U}_s$:

$$\begin{aligned}\widetilde{R^o} &= \widetilde{C} \times \widetilde{C'} \subseteq Þ\mathrm{R}, C' \neq C \wedge C' \sqsubseteq \mathfrak{U}_s \\ &= \mathop{R}_{k=1}^{|\mathfrak{O}|} \{(\widetilde{C}, \widetilde{C'_k}), \mu(\widetilde{R}_k) = \sigma)\}\end{aligned} \tag{6}$$

where $\widetilde{C'}$ is a fuzzy set of external concepts in $\mathfrak{U}_s$, and the membership $\mu(\widetilde{R}_k)$ is determined by the *conceptual equivalency* σ between the sets of fuzzy attributes from each fuzzy concepts, i.e.:

$$\sigma = \frac{|\widetilde{A} \cap \widetilde{A'}|}{|\widetilde{A} \cup \widetilde{A'}|} \tag{7}$$

- $\widetilde{Q}$ is a fuzzy set of *qualifications* that modifies the concept $\widetilde{C}$ by weights in (0, 1] as a special part of the external relations $\widetilde{R^o}$:

$$\widetilde{Q} = \{(q_1, \omega(q_1)), (q_2, \omega(q_2)), \ldots, (q_p, \omega(q_p)))\} \subseteq Þ\mathfrak{R} \tag{8}$$

where $\widetilde{Q}$ is initially empty when the concept is independent. However, it obtains qualified properties and weights when the fuzzy concept is modified by an adjective or adjective phrase, or it is comparatively evaluated with other fuzzy concepts.

In the fuzzy concept model, Eqs. 5 and 6 denote general internal and external relations, respectively. A concrete fuzzy relation in a specific fuzzy concept will be an instantiation of the general relations tailored by a given characteristic matrix on the Cartesian products.

As described in Definition 2, the important properties of a formal fuzzy concept are the fuzzy set of essential attributes as its *intension*; the fuzzy set of instantiated objects as its *extension*; and the adaptive capability to autonomously interrelate the concept to other concepts in an existing knowledge base in $\mathfrak{U}_s$.

Example 1. A fuzzy concept 'pen', $\widetilde{C}(pen)$, can be formally described according to Definition 2 as follows:

$$\begin{aligned}\widetilde{C}(pen) &\triangleq \widetilde{C}(\widetilde{A}, \widetilde{O}, \widetilde{R^i}, \widetilde{R^o}, \widetilde{Q}) \\ &= \widetilde{pen}((\widetilde{A}, \widetilde{\mu_C(A)}), (\widetilde{O}, \widetilde{\mu_C(O)}), \widetilde{R^{i'}}, \widetilde{R^{o'}}, \widetilde{Q}) \\ &= \begin{cases} \widetilde{A} = \{(a_1, \mu(a_1)), (a_2, \mu(a_2)), (a_3, \mu(a_3)), (a_4, \mu(a_4))\} \\ \quad = \{(writing\_tool,\ 1.0),\ (ink,\ 0.9),\ (nib,\ 0.9), \\ \quad\quad (ink\_container,\ 0.8)\} \\ \widetilde{O} = \{(o_1, \mu(o_1)), (o_2, \mu(o_2)), (o_3, \mu(o_3)), (o_4, \mu(o_4))\} \\ \quad = \{(ballpo\mathrm{int},\ 1.0),\ (fountain,\ 1.0), \\ \quad\quad (pencil,\ 0.9),\ (brush,\ 0.7)\} \\ \widetilde{R^i} = \widetilde{O} \times \widetilde{A} \\ \widetilde{R^o} = \widetilde{C} \times \widetilde{C'} \\ \widetilde{Q} = \varnothing \end{cases}\end{aligned}$$

ISBN: 978-1-61804-227-9

Example 2. A fuzzy concept 'man', $\widetilde{C}(man)$, can be formally described based on Definition 4 as follows:

$$\begin{aligned}\widetilde{C}(man) &\triangleq \widetilde{C}(\widetilde{A},\widetilde{O},\widetilde{R^i},\widetilde{R^o},\widetilde{Q}) \\ &= \widetilde{man}((\widetilde{A},\widetilde{\mu_C(A)}),(\widetilde{O},\widetilde{\mu_C(O)}),\widetilde{R^{i'}},\widetilde{R^{o'}},\widetilde{Q}) \\ &= \begin{cases}\widetilde{A} = \{(human\_being,\ 1.0),\ (male,\ 1.0),\ (adult,\ 0.9)\} \\ \widetilde{O} = \{(American,\ 1.0),\ (Australia,\ 1.0), \\ \qquad (business\_man,\ 1.0),\ ...\} \\ \widetilde{R^{i'}} = \widetilde{O}\times\widetilde{A} \\ \widetilde{R^{o'}} = \widetilde{C}\times\widetilde{C'} \\ \widetilde{Q} = \varnothing\end{cases}\end{aligned}$$

Applying the fuzzy concept model as a basic unit of semantic knowledge in $\mathfrak{U}_s$, the fuzzy semantics in natural languages can be expressed as a mapping from a fuzzy language entity to a determined fuzzy concept where its sets of fuzzy attributes, objects, relations, and qualifications are specified.

Definition 3. The *fuzzy semantics of an entity e*, $\widetilde{\Theta}(e)$, is an equivalent fuzzy concept $\widetilde{C}_e$ in $\mathfrak{U}_s$, i.e.:

$$\begin{aligned}\widetilde{\Theta}(e) &\triangleq \widetilde{\Theta}(e=\widetilde{C_e}) \\ &= \widetilde{C}_e(\widetilde{A_e},\widetilde{O_e},\widetilde{R^i_e},\widetilde{R^o_e},\widetilde{Q_e})\end{aligned} \tag{9}$$

where $\widetilde{C}_e$ is denoted according to the generic model of fuzzy concepts as given in Definition 2.

Example 3. The fuzzy semantics of a language entity 'pen', denoted by $\Theta(\widetilde{C}(pen))$, can be formally derived according to Definition 3 and Example 1 as follows:

$$\begin{aligned}\widetilde{\Theta}_e(pen) &\triangleq \widetilde{\Theta}(e=\widetilde{C}(pen)) \\ &= \widetilde{pen}((\widetilde{A},\widetilde{\mu_{\widetilde{C}}(A)}),(\widetilde{O},\widetilde{\mu_{\widetilde{C}}(O)}),\widetilde{R^i},\widetilde{R^o},\widetilde{Q}) \\ &= \begin{cases}\widetilde{A} = \{(writing\_tool,\ 1.0),\ (ink,\ 0.9),\ (nib,\ 0.9), \\ \qquad (ink\_container,\ 0.8)\} \\ \widetilde{O} = \{(ballpo\text{int},\ 1.0),\ (fountain,\ 1.0), \\ \qquad (pencil,\ 0.9),\ (brush,\ 0.7)\} \\ \widetilde{R^i} = \widetilde{O}\times\widetilde{A} \\ \widetilde{R^o} = \widetilde{C}\times\widetilde{C'} \\ \widetilde{Q} = \varnothing\end{cases}\end{aligned}$$

Example 4. Similarly, the fuzzy semantics of a language entity 'man', denoted by $\Theta(\widetilde{C}(man))$, can be formally derived based on Definition 3 and Example 2 as follows:

$$\begin{aligned}\widetilde{\Theta}_e(man) &\triangleq \widetilde{\Theta}(e=\widetilde{C}(man)) \\ &= \widetilde{man}((\widetilde{A},\widetilde{\mu_{\widetilde{C}}(A)}),(\widetilde{O},\widetilde{\mu_{\widetilde{C}}(O)}),\widetilde{R^{i'}},\widetilde{R^{o'}},\widetilde{Q}) \\ &= \begin{cases}\widetilde{A} = \{(human\_being,\ 1.0),\ (male,\ 1.0),\ (adult,\ 0.9)\} \\ \widetilde{O} = \{(American,\ 1.0),\ (Australia,\ 1.0), \\ \qquad (business\_man,\ 1.0),\ ...\} \\ \widetilde{R^{i'}} = \widetilde{O}\times\widetilde{A} \\ \widetilde{R^{o'}} = \widetilde{C}\times\widetilde{C'} \\ \widetilde{Q} = \varnothing\end{cases}\end{aligned}$$

Therefore, on the basis of Definition 3, fuzzy semantic analyses and comprehension in natural languages can be formally described as a deductive process from a fuzzy entity to a determined fuzzy concept.

Corollary 1. The *rule of semantic deduction* states that the semantics of a given fuzzy entity is comprehended in semantic analysis, if and only if its fuzzy semantics can be reduced onto a known fuzzy concept with determined membership and weight values.

III. Fuzzy Semantics of Modifiers on Concepts in Fuzzy Inferences

The semantics of fuzzy concepts is usually modified by an adjective or an adjective phrase in language expressions in order to fine tune its qualification such as degree, scope, quality, constraint, purpose, and etc. Therefore, the fuzzy semantics of fuzzy concepts as developed in Section 2 can be extended to deal with composite semantics of noun phrases modified by determiners and degree words [19, 27, 28, 31, 32, 34].

The modifier in cognitive linguistics is words or phrases that elaborate, limit, and qualify a noun or noun phrase in the categories of determiners, qualifiers, degrees, and negations [6, 28]. A fuzzy modifier can be represented as a fuzzy set with certain weights of memberships [31, 32, 34]. For instance, Zadeh considered the fuzzy effects of some special adverbs on adjectives such as 'very, very', 'very little', 'positive', and 'negative' in 1975, which were modeled as nonlinear exponential weights on the target adjectives [32]. However, the general semantics relations between a fuzzy linguistic entity (noun) and its fuzzy modifier (adverb-adjective phrase) are yet to be studied.

Definition 4. A *fuzzy modifier* $\widetilde{\tau}$ is a special fuzzy set that represents an adjective or adjective phrase in natural languages where its memberships are replaced by intentional weights of the modifier, $\omega_{\widetilde{\tau}}(\tau_k),\ 1 \le k \le z,$ and z is a constant, i.e.:

$$\begin{aligned}\tilde{\tau} &= \{\mathop{R}_{k=1}^{z}(\tau_k, \omega_{\tilde{\tau}}(\tau_k))\},\ \omega_{\tilde{\tau}}(\tau_k) \in (0,1] \\ &= \{(\tau_1,\omega(\tau_1)),(\tau_2,\omega(\tau_2)),\ldots,(\tau_z,\omega(\tau_z)\}\end{aligned} \tag{10}$$

where the weights of $\tilde{\tau}$ is normalized in the domain $(0,1]$.

Example 5. A fuzzy modifier 'good' on the *quality* of a fuzzy entity can be formally described as a fuzzy set $\tilde{\tau}(good)$ according to Definition 4 as follows:

$$\begin{aligned}\tilde{\tau}(good) &= \{\mathop{R}_{k=1}^{4}(\tau_k(quality),\ \mu_{\tilde{\tau}}(\tau_k))\} \\ &= \{(neutral,\ 0.1),\ (ok,\ 0.6), \\ &\quad (excellent,\ 0.8),\ (perfect,\ 1.0)\}\end{aligned} \tag{11}$$

Example 6. A fuzzy modifier 'old' on the fuzzy entity *ages* can be formally described as a fuzzy set $\tilde{\tau}(old)$ as follows:

$$\begin{aligned}\tilde{\tau}(old) &= \{\mathop{R}_{k=1}^{6}(\tau_k(age),\ \omega_{\tilde{\tau}}(\tau_k))\} \\ &= \{([\text{1-20}],\ 0),\ ([\text{21-30}],\ 0.1),\ ([\text{31-50}],\ 0.4), \\ &\quad ([\text{51-65}],\ 0.7),\ ([\text{66-80}],\ 0.9),\ ([>80],\ 1.0)\}\end{aligned} \tag{12}$$

Definition 5. A *fuzzy qualifier* $\tilde{\delta}$ is a *special* fuzzy set of degree adverbs or adverb phrases to modify $\tilde{\tau}$ in natural languages where their memberships are replaced by *intentional weights* of degree and extends, $\omega_{\tilde{\delta}_l}(\delta_l),\ 1 \le l \le q,\ \text{and}\ q\ \text{is a constant}$, i.e.:

$$\begin{aligned}\tilde{\delta} &= \{\mathop{R}_{l=1}^{p}(\delta_l, \omega_{\tilde{\delta}}(\delta_l))\},\ \omega_{\tilde{\delta}}(\delta_l) \in \pm(0,3] \\ &= \{(\delta_1,\omega(\tau_1)),(\delta_2,\omega(\delta_2)),\ldots,(\delta_p,\omega(\delta_p)\}\end{aligned} \tag{13}$$

where the weights of $\tilde{\delta}$ is constrained in the domain $\pm[1,3]$ corresponding to the *neutral* (1), *comparative* (2), and *superlative* (3) degrees of adverbs and adjectives in natural languages.

Example 7. A typical fuzzy set of qualifiers, $\tilde{\delta}$, can be described according to Definition 5 as follows:

$$\begin{aligned}\tilde{\delta} = \{&(definitely\_not,\ \text{-}3.0),\ (imperfectly,\ \text{-}1.5), \\ &(neutral\_negative,\ \text{-}1.0),\ (somewhat,\ 0.5), \\ &(fairly,\ 1.2),\ (quite,\ 1.5),\ (excellently,\ 2.0), \\ &(extremely,\ 3.0)\},\ \omega_{\tilde{\delta}}(\delta) \in \pm(0,3]\end{aligned} \tag{14}$$

Definition 6. A *composite fuzzy modifier* $\widetilde{\delta\tau}$ is a product of a fuzzy qualifier $\tilde{\delta}$ and a fuzzy modifier $\tilde{\tau}$. The value of the composite modifiers is determined by the product of their weights, i.e.:

$$\begin{aligned}\Theta(\tilde{\delta} \bullet \tilde{\tau}) &\triangleq \Theta(\widetilde{\delta(y) \bullet \tau(x)}) \\ &= \omega_{\tilde{\delta}}(y) \bullet \omega_{\tilde{\tau}}(x),\ 0 < \omega_{\tilde{\tau}}(x) \le 1, \\ &\qquad \text{-}3 \le \omega_{\tilde{\delta}}(y) \le 3,\ \omega_{\tilde{\delta}}(y) \ne 0\end{aligned} \tag{15}$$

where the *combined domain* of composite modifiers is $\pm(0,3]$ in order to be consistent to the modifiers in real-world languages.

In case a weight of the fuzzy qualifiers is less than zero, the composite modifier represents a negative intention. For instance, $\tilde{\delta} \bullet \tilde{\tau} = \tilde{\delta}(neutral\_negative) \bullet \tilde{\tau}(good)$ implies a weight of qualification in the semantics as $\omega(\tilde{\delta}(neutral\_negative)) \bullet \omega(\tilde{\tau}(good)) = -1 \bullet 0.6 = -0.6.$

On the basis of the formal semantics of fuzzy modifiers $\tilde{\tau}$, qualifiers $\tilde{\delta}$, and composite modifiers $\widetilde{\tau'} = \widetilde{\delta\tau}$, the composite fuzzy semantics of language entities modified by $\widetilde{\delta\tau}$ can be quantitatively expressed.

Definition 7. The *composite fuzzy semantic* of a fuzzy concept $\tilde{C}$ qualified by a *fuzzy* modifier $\tilde{\tau}$, qualifier $\tilde{\delta}$, and/or a composite fuzzy modifier $\widetilde{\tau'} = \widetilde{\delta\tau}$, denoted by $\tilde{\Theta}(\widetilde{C'}) = \tilde{\Theta}(\widetilde{\tau'} \bullet \tilde{C})$, is a complex semantics of the fuzzy concept $\tilde{C}$ qualified by a certain weight of the composite modifier, i.e.:

$$\begin{aligned}\tilde{\Theta}(\widetilde{\tau'} \bullet \tilde{C}) &\triangleq \tilde{\Theta}(\tilde{C} \mid \tilde{Q} = \widetilde{\delta\tau}) \\ &= \tilde{C}(\tilde{A},\tilde{O},\widetilde{R^i},\widetilde{R^o},\tilde{Q} \mid \tilde{Q} = \widetilde{\delta\tau}) \\ &= \tilde{C}((\tilde{A},\widetilde{\mu_{\tilde{A}}(A)}),(\tilde{O},\widetilde{\mu_{\tilde{O}}(O)}),\widetilde{R^i},\widetilde{R^o},(\tilde{Q} = \widetilde{\delta\tau})) \\ &= \tilde{C}(\tilde{A},\tilde{O},\widetilde{R^i},\widetilde{R^o},\tilde{Q})\end{aligned} \tag{16}$$

where the fuzzy set of composite modifiers imposes a specific set of weights of intentional qualifications $\tilde{Q}$ in the modified semantics of the target fuzzy concept, and $\tilde{\delta} = 1$ if it is absent.

Example 8. Given a fuzzy concept $\tilde{C}(pen)$ as obtained in Example 1, the composite semantics $\tilde{\Theta}(\tilde{\tau} \bullet \tilde{C}) = \tilde{\Theta}(excellent\_pen)$ qualified by the fuzzy modifier $\tilde{\tau}(good)$ can be determined according to Definition 7, i.e.:

$$\begin{aligned}\tilde{\Theta}(\widetilde{\delta\tau} \bullet \tilde{C}) &= \tilde{\Theta}(\tilde{\tau}(good) \bullet \widetilde{pen}) \\ &= \widetilde{excellent\_pen}((\tilde{A},\widetilde{\mu_{\tilde{A}}(A)}),(\tilde{O},\widetilde{\mu_{\tilde{O}}(O)}),\widetilde{R^i},\widetilde{R^o}, \\ &\qquad (\tilde{Q} = \tau_0 \mid \tau_0 = \tau(excellent) = 0.8)) \\ &= \begin{cases} \tilde{A} = \{(writing,\ 1.0),\ (ink,\ 0.9),\ (nib,\ 0.9), \\ \quad (ink\_container,\ 0.8)\} \\ \tilde{O} = \{(ballpoint,\ 1.0),\ (fountain,\ 1.0),\ (pencil,\ 0.9), \\ \quad (brush,\ 0.7)\} \\ \widetilde{R^i} = \tilde{O} \times \tilde{A} \\ \widetilde{R^o} = \tilde{C} \times \widetilde{C'} \\ \tilde{Q}(excellent) = 0.8 \end{cases}\end{aligned}$$

Example 9. The fuzzy semantics of $\widetilde{C}(excellent\_pen)$ obtained in Example 8 may be further modified by a qualifier $\tilde{\delta}(extremely)$ that results in $\widetilde{C}(extremely\_excellent\_pen)$ as follows:

$$\begin{aligned}
&\widetilde{\Theta}(\widetilde{\delta\tau} \bullet \widetilde{C}) = \widetilde{\Theta}(\tilde{\delta}(extremely) \bullet \tilde{\tau}(good) \bullet \widetilde{pen}) \\
&= \widetilde{extremely_excellent_pen}((\widetilde{A}, \mu_{\widetilde{A}}(A)), (\widetilde{O}, \mu_{\widetilde{O}}(O)), \widetilde{R^i}, \widetilde{R^o}, \\
&\quad (\widetilde{Q} = \tau'_0 \mid \tau'_0 = \tilde{\delta}(extremely) \bullet \tau(excellent) = 3.0 \bullet 0.8)) \\
&= \begin{cases}
\widetilde{A} = \{(writing_tool,\ 1.0),\ (ink,\ 0.9),\ (nib,\ 0.9), \\
\qquad (ink_container,\ 0.8)\} \\
\widetilde{O} = \{(ballpoint,\ 1.0),\ (fountain,\ 1.0),\ (pencil,\ 0.9), \\
\qquad (brush,\ 0.7)\} \\
\widetilde{R^i} = \widetilde{O} \times \widetilde{A} \\
\widetilde{R^o} = \widetilde{C} \times \widetilde{C}' \\
\widetilde{Q}(extremely_excellent) = 2.4
\end{cases}
\end{aligned}$$

Example 10. Given a fuzzy concept $\widetilde{C}(man)$ as described in Example 2, the composite semantics $\widetilde{\Theta}(\tilde{\tau} \bullet \widetilde{C}) = \widetilde{\Theta}(old\_man)$ qualified by the fuzzy modifier $\tilde{\tau}$ (*old*) can be determined according to Definition 7 as follows:

$$\begin{aligned}
\widetilde{\Theta}(\tilde{\tau} \bullet \widetilde{C}) &= \widetilde{\Theta}(\tilde{\tau}(old) \bullet \widetilde{man}),\ \tau_0 = \tau(60) = 0.7 \\
&= \widetilde{old_man}((\widetilde{A}, \mu_{\widetilde{A}}(A)), (\widetilde{O}, \mu_{\widetilde{O}}(O)), \widetilde{R^i}, \widetilde{R^o}, (\widetilde{Q} = \tau_0)) \\
&= \begin{cases}
\widetilde{A} = \{(human_being,\ 1.0),\ (male,\ 1.0), \\
\qquad (adult,\ 0.9)\} \\
\widetilde{O} = \{(American,\ 1.0),\ (Australia,\ 1.0), \\
\qquad (business_man,\ 1.9),\ ...\} \\
\widetilde{R^i} = \widetilde{O} \times \widetilde{A} \\
\widetilde{R^o} = \widetilde{old_man} \times \widetilde{C}' \\
\widetilde{Q}(old) = 0.7
\end{cases}
\end{aligned}$$

Example 10 indicates that, against the fuzzy qualifier $\tilde{\tau}(old)$, a man in the age of 60 is 0.7 (*quite likely*) as an old man. Similarly, other instantiations modified by the qualifiers may denote that a man in the age of 25 is 0.1 (*unlikely*) as an old man; and a man in the age of 85 is 1.0 (*definitely*) as an old man.

Example 11. The fuzzy semantics of $\widetilde{C}(old\_man)$ obtained in Example 10 may be further modified by a qualifier $\tilde{\delta}(quite)$ that results in $\widetilde{C}(a\_quite\_old\_man)$ as follows:

$$\begin{aligned}
\widetilde{\Theta}(\widetilde{\delta\tau} \bullet \widetilde{C}) &= \widetilde{\Theta}(\tilde{\delta}(quite) \bullet \tilde{\tau}(old) \bullet \widetilde{man}) \\
&= \widetilde{a_quite_old_man}((\widetilde{A}, \mu_{\widetilde{A}}(A)), (\widetilde{O}, \mu_{\widetilde{O}}(O)), \widetilde{R^i}, \widetilde{R^o}, \\
&\qquad (\widetilde{Q} = \tau'_0 \mid \tau'_0 = \tilde{\delta}(quite) \bullet \tau(60) = 1.5 \bullet 0.7)) \\
&= \begin{cases}
\widetilde{A} = \{(human_being,\ 1.0),\ (male,\ 1.0),\ (adult,\ 0.9)\} \\
\widetilde{O} = \{(American,\ 1.0),\ (Australia,\ 1.0), \\
\qquad (business_man,\ 1.0),\ ...\} \\
\widetilde{R^i} = \widetilde{O} \times \widetilde{A} \\
\widetilde{R^o} = \widetilde{old_man} \times \widetilde{C}' \\
\widetilde{Q}(quite_old) = 1.05
\end{cases}
\end{aligned}$$

The fuzzy nature of language semantics and their comprehension is formally explained by the mathematical models of fuzzy concepts and fuzzy semantics qualified by fuzzy modifiers. Based on the formal theory, fuzzy semantic inferences can be rigorously manipulated to deal with fuzzy degrees of matters, uncertainties, vague semantics, and fuzzy causality in cognitive linguistics and fuzzy systems. This work enables cognitive machines, cognitive robots, and fuzzy system to mimic the human intelligent ability and the cognitive processes in cognitive linguistics, fuzzy inferences, cognitive computing, and computational intelligence.

IV. Conclusion

The mathematical models of fuzzy concepts and fuzzy semantics have provided a formal explanation for the fuzzy nature of human language processing and real-time semantics interpretation. It has been identified that the basic unit of linguistic entities that carries unique and unambiguous semantics is a fuzzy concept, which can be modeled as a fuzzy hyperstructure encompassing fuzzy sets of attributes, objects, relations, and qualifications. Complex fuzzy concepts in natural languages have been modeled as a composite fuzzy concept where fuzzy qualifiers are involved to modify the fuzzy concept. As a result, the fuzzy semantics of composite fuzzy concepts has been denoted as algebraic operations on the fuzzy qualification of the target fuzzy concept. This work has demonstrated that fuzzy semantic comprehension is a deductive process, where complex fuzzy semantics can be formally expressed by algebraic operations on elementary ones with fuzzy modifiers. The denotational mathematical structure of fuzzy concepts and fuzzy semantics not only reveals the fuzzy properties of human semantic comprehension, but also enables cognitive machines and fuzzy systems to mimic the human fuzzy inference mechanisms in cognitive linguistics, fuzzy systems, cognitive computing, and computational intelligence.

ISBN: 978-1-61804-227-9

Acknowledgment

A number of notions in this work have been inspired by Prof. Lotfi A. Zadeh during the author's sabbatical leave at BISC, UC Berkeley as a visiting professor. I am grateful to Dr. Zadeh for his vision, insight, support, and enjoyable discussions. The author would like to thank the anonymous reviewers for their valuable comments on the previous version of this paper.

References

[1] E. A. Bender, *Mathematical Methods in Artificial Intelligence*, Los Alamitos, CA: IEEE CS Press, 1996.
[2] BISC, *Internal Communications*, University of California, Berkeley, CA., 2012.
[3] N. Chomsky, "Three Models for the Description of Languages," *I.R.E. Trans. Information Theory*, vol. 2, no.3, pp. 113-124, 1956.
[4] N. Chomsky, *Aspects of the Theory of Syntax*, Cambridge, MA: MIT Press, 1965.
[5] V. Evans and M. Green, *Cognitive Linguistics: An Introduction*, Edinburgh: Edinburgh University Press, 2006.
[6] H. Eugene, "Cognitive Linguistics in the Redwoods: The Expansion of a New Paradigm in Linguistics," in *Cognitive Linguistics Research*, 6. Berlin, 1996.
[7] R. Jackendoff, *Foundations of Language: Brain, Meaning, Grammar, and Evolution*, UK: Oxford Univ. Press, 2002.
[8] E. Keenan, ed., *Formal Semantics of Natural Language*, UK: Cambridge University Press, 1975.
[9] R. Montague, "English as a Formal Language," in R. H. Thomason ed., *Formal Philosophy; Selected Papers of Richard Montague*, London: Yale University Press, pp. 188-221, 1974.
[10] W. O'Grady and J. Archibald, *Contemporary Linguistic Analysis: An Introduction*, 4th ed., Toronto: Pearson Education Canada Inc., 2000.
[11] B. H. Partee, A. ter Meulen, and R. Wall, Eds., *Mathematical Methods in Linguistics*, Dordrecht: Kluwer, 1990.
[12] S. Pullman, *Computational Linguistics*, Cambridge, UK: Cambridge University Press, 1997.
[13] T. J. Ross, *Fuzzy Logic with Engineering Applications*, McGraw-Hill Co., 1995.
[14] J. I. Saeed, Semantics, 3rd ed., Oxford, UK: Wiley-BlackWell, 2009.
[15] R. J. Sternberg, *In Search of the Human Mind*, 2nd ed., NY: Harcourt Brace & Co., 1998.
[16] Y. Wang, "On Cognitive Informatics," J. *Brain and Mind*, vol. 4, no. 3, pp. 151-167, 2003.
[17] Y. Wang, "The Cognitive Processes of Formal Inferences," *International J. Cognitive Informatics and Natural Intelligence*, vol.1, no. 4, pp. 75-86, 2007.
[18] Y. Wang, "On Contemporary Denotational Mathematics for Computational Intelligence," *Trans. Computational Science,* vol. 2, *pp.* 6-29, 2008.
[19] Y. Wang, "On Concept Algebra: A Denotational Mathematical Structure for Knowledge and Software Modeling," *International J. Cognitive Informatics and Natural Intelligence*, vol. 2, no. 2, pp. 1-19, 2008.
[20] Y. Wang, "On Cognitive Computing," *International J. Software Science and Computational Intelligence*, vol. 1, no. 3, pp. 1-15, 2009.
[21] Y. Wang, "Cognitive Robots: A Reference Model towards Intelligent Authentication," *IEEE Robotics and Automation*, vol. 17, no. 4, pp. 54-62, 2010.
[22] Y. Wang, "On Formal and Cognitive Semantics for Semantic Computing," *International J. Semantic Computing*, vol. 4, no. 2, pp. 203–237, 2010.
[23] Y. Wang, "Inference Algebra (IA): A Denotational Mathematics for Cognitive Computing and Machine Reasoning (I)," *International J. Cognitive Informatics and Natural Intelligence*, vol. 5, no.4, pp. 62-83, 2011.
[24] Y. Wang, "On Cognitive Models of Causal Inferences and Causation Networks," *International J. Software Science and Computational Intelligence*, vol. 3, no. 1, pp. 50-60, 2011.
[25] Y. Wang, "In Search of Denotational Mathematics: Novel Mathematical Means for Contemporary Intelligence, Brain, and Knowledge Sciences," *J. Advanced Mathematics and Applications*, vol. 1, no. 1, pp. 4-25, 2012.
[26] Y. Wang, "Formal Rules for Fussy Causal Analyses and Fuzzy Inferences," *International J. Software Science and Computational Intelligence*, vol. 4, no. 4, pp. 70-86, 2012.
[27] Y. Wang, "A Semantic Algebra for Cognitive Linguistics and Cognitive Computing," in *Proc. 12th IEEE International Conference on Cognitive Informatics and Cognitive Computing* (ICCI*CC 2013), New York, IEEE CS Press, pp. 17-25, July 2013..
[28] Y. Wang and R. C. Berwick, "Towards a Formal Framework of Cognitive Linguistics," *J. Advanced Mathematics and Applications*, vol. 1, no. 2, pp. 250-263, 2012.
[29] R. A. Wilson and F. C. Kei, Eds., *The MIT Encyclopedia of the Cognitive Sciences,* Cambridge, MA: MIT Press, 2001.
[30] L. A. Zadeh, "Fuzzy Sets," *Information and Control*, vol. 8, pp. 338-353, 1965.
[31] L. A. Zadeh, "Quantitative Fuzzy Semantics," *Information Science*, vol. 3, pp. 159-176, 1971.
[32] L. A. Zadeh, "The Concept of a Linguistic Variable and its Application to Approximate Reasoning," Parts 1-3, *Information Science*, vol. 8, pp. 199-249, 301-357, and vol. 9, pp. 43-80, 1975.
[33] L. A. Zadeh, "Fuzzy Logic and Approximate Reasoning," *Syntheses*, vol. 30, pp. 407-428, 1975.
[34] L. A. Zadeh, "A Computational Approach to Fuzzy Quantifiers in Natural Languages," *Computers and Mathematics*, vol. 9, pp. 149-181, 1983.
[35] L. A. Zadeh, "From Computing with Numbers to Computing with Words – from Manipulation of Measurements to Manipulation of Perception," *IEEE Trans. on Circuits and Systems*, vol. 45, pp. 105-119, 1999.
[36] L. A. Zadeh, "Precisiated Natural Language (PNL)," *AI Magazine*, vol. 25, no. 3, pp. 74-91, 2004.
[37] L. A. Zadeh, "Is There a Need for Fuzzy Logic?" *Information Sciences*, vol. 178, pp. 2751-2779, 2008.

On Statistical Limit Points in a Fuzzy Valued Metric Space

S. Aytar, U. Yamancı, and M. Gürdal

***Abstract*—**We introduce the concepts statistical cluster and statistical limit points of a sequence of fuzzy numbers in a fuzzy valued metric space. Then we obtain some inclusion relations between the sets of limit points, statistical limit points and statistical cluster points for a sequence of fuzzy numbers.

***Keywords*—** Statistical convergence; Statistical limit point; Fuzzy valued metric space.

I. Introduction and Background

OVER the past few years the theory of convergence of a sequence of fuzzy numbers has been studied by many authors [1-3,8,11,19]. The first steps towards constructing such convergence theories go back to Matloka's [13] and Kaleva's [12] works. To this end, they used the supremum metric that gives a real (crisp) value for the distance between two fuzzy numbers. On the other hand, via positive fuzzy numbers, it is also possible to define a fuzzy (non-crisp) distance between two fuzzy numbers (as is exemplified by Guangquan [9]), because it is more natural that the distance between two fuzzy numbers is a fuzzy number rather than this distance is a real number.

In this paper, we introduce the concept of statistical convergence of a sequence of fuzzy numbers in a fuzzy metric space which defined by Guangquan [9,10]. Then we compare this definition with the definition of statistical convergence with respect to the supremum metric. We note that this convergence should not be perceived as a generalization of ordinary statistical convergence (see Example 1). Moreover we define the concepts statistical cluster and statistical limit points of a sequence of fuzzy numbers in this fuzzy valued metric space. Finally we obtain some inclusion relations between the sets of limit points, statistical limit points and statistical cluster points for a sequence of fuzzy numbers.

S. Aytar is with the Suleyman Demirel University, Department of Mathematics, 32260, Isparta, TURKEY (corresponding author to provide phone: +90-246-2114120; fax: +90-246-2371106; e-mail: salihaytar@sdu.edu.tr).

U. Yamancı is with the Suleyman Demirel University, Department of Mathematics, 32260, Isparta, TURKEY (corresponding author to provide phone: +90-246-2114310; fax: +90-246-2371106; e-mail: ulasyamanci@sdu.edu.tr).

M. Gürdal, is with the Suleyman Demirel University, Department of Mathematics, 32260, Isparta, TURKEY (corresponding author to provide phone: +90-246-2114101; fax: +90-246-2371106; e-mail: gurdalmehmet@sdu.edu.tr).

Now we recall some definitions and notations that will be used frequently (see [4-7,10,12,13-18] for more details).

Given an interval A, we denote its endpoints by $\underline{A}$ and $\overline{A}$. We denote by D the set of all closed intervals on the real line $\mathbb{R}$. That is,

$$D := \{A \subset \mathbb{R} : \quad A = [\underline{A}, \overline{A}]\}.$$

For $A, B \in D$ define

$$A \le B \quad \text{iff} \quad \underline{A} \le \underline{B} \quad \text{and} \quad \overline{A} \le \overline{B},$$
$$d(A,B) := \max(|\underline{A} - \underline{B}|, |\overline{A} - \overline{B}|).$$

It is easy to see that d defines a metric (Hausdorff metric) on D and (D,d) is a complete metric space. Also "$\le$" is a partial order on D.

A fuzzy number is a function X from $\mathbb{R}$ to $[0,1]$, satisfying

- X is normal, i.e., there exists $x_0 \in \mathbb{R}$ such that $X(x_0) = 1$;
- X is fuzzy convex, i.e., for any $x, y \in \mathbb{R}$ and $\lambda \in [0,1], X(\lambda x + (1-\lambda) y) \ge \min\{X(x), X(y)\}$;
- X is upper semi-continuous;
- the closure of the set $\{x \in \mathbb{R} : X(x) > 0\}$, denoted by X^0, is compact.

These properties imply that for each $\alpha \in (0,1]$, the $\alpha-$level set $X^\alpha := \{x \in \mathbb{R} : X(x) \ge \alpha\} = \left[\underline{X}^\alpha, \overline{X}^\alpha\right]$ is a nonempty compact convex subset of $\mathbb{R}$, as the support X^0 : $X^0 = \lim_{\alpha \to 0^+} X^\alpha$ is We denote the set of all fuzzy numbers by $F(\mathbb{R})$. Note that the function a_1 defined by

$$a_1(x) := \begin{cases} 1 & ,\text{if } x = a \\ 0 & ,\text{otherwise} \end{cases},$$

where $a \in \mathbb{R}$, is a fuzzy number. By the *decomposition theorem* of fuzzy sets, we have

$$X = \sup_{\alpha \in [0,1]} \alpha \chi_{\left[\underline{X}^\alpha, \overline{X}^\alpha\right]}$$

for every $X \in \mathbf{F}(\mathrm{R})$, where each $\chi_{\left[\underline{X}^{\alpha},\overline{X}^{\alpha}\right]}$ denotes the characteristic function of the subinterval $\left[\underline{X}^{\alpha},\overline{X}^{\alpha}\right]$.

Now we recall the *partial order relation* on the set of fuzzy numbers. For $X,Y \in \mathbf{F}(\mathrm{R})$, we write $X \le Y$, if for every $\alpha \in [0,1]$, the inequality

$$X^{\alpha} \le Y^{\alpha}$$

holds. We write $X < Y$, if $X \le Y$, and there exists an $\alpha_0 \in [0,1]$ such that

$$\underline{X}^{\alpha_0} < \underline{Y}^{\alpha_0} \text{ or } \overline{X}^{\alpha_0} < \overline{Y}^{\alpha_0}.$$

If $X \le Y$ and $Y \le X$, then $X = Y$. Two fuzzy numbers X and Y are said to be *incomparable* and denoted by $X \not\sim Y$, if neither $X \le Y$ nor $Y \le X$ holds. When $X \ge Y$ or $X \not\sim Y$, then we can write $X \not< Y$.

Now let us briefly review the operations of *summation* and *subtraction* on the set of fuzzy numbers. For $X,Y,Z \in \mathbf{F}(\mathrm{R})$, the fuzzy number Z is called the *sum of* X *and* Y, and we write $Z = X + Y$, if $Z^{\alpha} = \left[\underline{Z}^{\alpha},\overline{Z}^{\alpha}\right] := X^{\alpha} + Y^{\alpha}$ for every $\alpha \in [0,1]$. Similarly, we write $Z = X - Y$, if $Z^{\alpha} = \left[\underline{Z}^{\alpha},\overline{Z}^{\alpha}\right] := X^{\alpha} - Y^{\alpha}$ for every $\alpha \in [0,1]$.

We define the set of *positive fuzzy numbers* by

$$\mathbf{F}^{+}(\mathrm{R}) := \left\{ X \in \mathbf{F}(\mathrm{R}) : X \ge 0_1 \text{ and } \overline{X}^{1} > 0 \right\}.$$

The map $d_M : \mathbf{F}(\mathrm{R}) \times \mathbf{F}(\mathrm{R}) \to R^{+} \cup \{0\}$ defined as

$$d_M(X,Y) := \sup_{\alpha\in[0,1]} d\left(X^{\alpha},Y^{\alpha}\right)$$

is called the *supremum metric* on $\mathbf{F}(\mathrm{R})$.

A sequence $X = \{X_n\}$ of fuzzy numbers is said to be *convergent* to the fuzzy number X_0, written as $\lim X_n = X_0$, if for every $\varepsilon > 0$ there exists a positive integer $n_0 = n_0(\varepsilon)$ such that

$$d_M(X_n, X_0) < \varepsilon \quad \text{for every } n > n_0.$$

A fuzzy number λ is called a *limit point* of the sequence $X = \{X_n\}$ of fuzzy numbers provided that there is a subsequence of X that converges to λ. We will denote the set of all limit points of $X = \{X_n\}$ by L_X.

Guangquan [9] introduced the concept of fuzzy distance between two fuzzy numbers as in Definition 1, and thus presented a concrete fuzzy metric in (1.1), which is very similar to an ordinary metric.

Definition 1. *A map* $\rho : \mathbf{F}(\mathrm{R}) \times \mathbf{F}(\mathrm{R}) \to \mathbf{F}(\mathrm{R})$ *is called a fuzzy metric on* $\mathbf{F}(\mathrm{R})$ *provided that the conditions*

(i) $\rho(X,Y) \ge 0_1$,

(ii) $\rho(X,Y) = 0_1$ *if and only if* $X = Y$,

(iii) $\rho(X,Y) = \rho(Y,X)$,

(iv) $\rho(X,Y) \le \rho(X,Z) + \rho(Z,Y)$ *are satisfied for all* $X,Y,Z \in \mathbf{F}(\mathrm{R})$.

If ρ is a fuzzy metric on the set of fuzzy numbers, then we call the triple $(\mathrm{R}, \mathbf{F}(\mathrm{R}), \rho)$ a *fuzzy metric space*. Guangquan [9] presented an example of a fuzzy metric space via the function d_G defined by

$$d_G(X,Y) := \sup_{\alpha\in[0,1]} \alpha \chi_{\left[\left|\underline{X}^{1}-\underline{Y}^{1}\right|,\ \sup_{t\in[\alpha,1]} d\left(X^{t},Y^{t}\right)\right]}. \tag{1.1}$$

Here the map d_G satisfies the conditions (i)-(iv) in Definition 1.

Remark 1. *Let*

$$\mathbf{B}_F := \left\{ K(X,P) : X \in \mathbf{F}(\mathrm{R}),\ P \in \mathbf{F}^{+}(\mathrm{R}) \right\} \subset \mathbf{P}(\mathbf{F}(\mathrm{R}))$$

where $\mathbf{P}(\mathbf{F}(\mathrm{R}))$ *is the power set of* $\mathbf{F}(\mathrm{R})$ *and* $K(X,P) := \left\{ Z \in \mathbf{F}(\mathrm{R}) : d_G(X,Z) < P,\ P \in \mathbf{F}^{+}(\mathrm{R}) \right\}$. *Then the set* $\mathbf{B}_F$ *forms a basis of a natural topology on* $\mathbf{F}(\mathrm{R})$, *denoted by* τ_F. *Thus the pair* $(\mathbf{F}(\mathrm{R}), \tau_F)$ *is a topological space.*

Now we investigate the properties of the convergence of a sequence in this topological space. Since this convergence is in the topology τ_F, we will denote it by $\tau_F -$convergence.

Definition 2 ($\tau_F -$convergence)**.** *Let* $X = \{X_n\} \subset \mathbf{F}(\mathrm{R})$ *and* $X_0 \in \mathbf{F}(\mathrm{R})$. *Then* $\{X_n\}$ *is* $\tau_F -$*convergent to* X_0 *and we denote this by*

$$\tau_F - \lim X_n = X_0 \text{ or } \{X_n\} \xrightarrow{\tau_F} X_0 \ (n \to \infty),$$

provided that for any $P \in \mathbf{F}^{+}(\mathrm{R})$ *there exists an* $n_0 = n_0(P) \in \mathrm{N}$ *such that*

$$d_G(X_n, X_0) < P \quad \text{as } n > n_0.$$

Definition 3 ($\tau_F -$limit point)**.** *A fuzzy number* λ *is a* $\tau_F -$*limit point of the sequence* $X = \{X_n\}$ *of fuzzy numbers provided that there is a subsequence of* X *that* $\tau_F -$*converges to* λ. *We denote the set of all* $\tau_F -$*limit points of* $X = \{X_n\}$ *by* $L_X^{\tau_F}$.

Let K be a subset of the set N of positive integers and let us denote the set $\{k \in K : k \le n\}$ by K_n. Then the *natural*

ISBN: 978-1-61804-227-9

density of K is defined by $\delta(K) := \lim_{n\to\infty} \frac{|K_n|}{n}$, where $|K_n|$ denotes the number of elements in K_n. Clearly, a finite subset has natural density zero and we have $\delta(K^c) = 1-\delta(K)$ whenever $\delta(K)$ exists, where $K^c := \mathsf{N} \setminus K$ is the complement of $K \subset \mathsf{N}$. If $K_1 \subseteq K_2$, then $\delta(K_1) \le \delta(K_2)$. In addition, $\delta(K) \neq 0$ means that $\overline{\delta}(K) := \limsup_{n\to\infty} \frac{|K_n|}{n} > 0.$

A sequence $X = \{X_n\}$ of fuzzy numbers is said to be statistically convergent to the fuzzy number X_0, written as $st-\lim X_n = X_0$, if the set

$$\{n \in \mathsf{N} \,:\, d_M(X_n, X_0) \ge \varepsilon\}$$

has natural density zero for every $\varepsilon > 0$.

Definition 4 (Nonthin subsequence)**.** *If $\{X_{n(j)}\}$ is a subsequence of $X = \{X_n\}$ and $K := \{n(j) \in \mathsf{N} \,:\, j \in \mathsf{N}\}$ then we abbreviate $\{X_{n(j)}\}$ by $\{X\}_K$ which in case $\delta(\{K\}) = 0$ is called a subsequence of density zero or thin subsequence. On the other hand, $\{X\}_K$ is a nonthin subsequence of X if K does not have density zero.*

Definition 5 (Statistical limit point)**.** *The fuzzy number ν is called statistical limit point of the sequence $X = \{X_n\}$ of fuzzy numbers provided that there is a nonthin subsequence of X that converges to ν. Let Λ_X denote the set of statistical limit points of the sequence X.*

Definition 6 (Statistical cluster point)**.** *The fuzzy number μ is called statistical cluster point of the sequence $X = \{X_n\}$ of fuzzy numbers provided that*

$$\overline{\delta}(\{n \in \mathsf{N} \,:\, d_M(X_n, \mu) < \varepsilon\}) > 0$$

for every $\varepsilon > 0$. Let Γ_X denote the set of statistical cluster point of the sequence X.

II. τ_F – STATISTICAL CONVERGENCE

Definition 7 (τ_F – statistical convergence)**.** *Let $X = \{X_n\}$ be a sequence of fuzzy numbers and X_0 be a fuzzy number. The sequence X is said to be τ_F – statistically convergent to X_0, we denote this by*

$$\tau_F\text{ - st - }\lim X_n = X_0,$$

provided that for any $P \in \mathsf{F}^+(\mathsf{R})$, the set

$$\{n \in \mathsf{N} \,:\, d_G(X_n, X_0) \not\succ P\}$$

has natural density zero.

It is clear that if a sequence is τ_F – convergent then it is τ_F – statistically convergent to the same fuzzy number. But the converse of this claim does not hold in general.

Example 1. *It is obvious that the sequence $X = \{X_n\}$ defined by*

$$X_n(x) := \begin{cases} \left.\begin{cases} 0 & \text{, if } x \in (-\infty, n-1) \cup (n+1, \infty) \\ x-(n-1) & \text{, if } x \in [n-1, n] \\ -x+(n+1) & \text{, otherwise} \end{cases}\right\} & \text{, if } n = k^2 \ (k \in \mathsf{N}) \\ \left.\begin{cases} 1-\frac{nx}{n+1} & \text{, if } x \in [0, 1+\frac{1}{n}] \\ 0 & \text{, otherwise} \end{cases}\right\} & \text{, otherwise} \end{cases}$$

is τ_F – statistically convergent to the fuzzy number

$$X_0(x) := \begin{cases} 1-x & \text{, if } x \in [0,1] \\ 0 & \text{, otherwise} \end{cases}.$$

On the other hand, since the set $\{n \in \mathsf{N} \,:\, d_G(X_n, X_0) \not\succ P\}$ has infinitely many elements for every $P \in \mathsf{F}^+(\mathsf{R})$, we say that this sequence is not τ_F – convergent.

Theorem 1. *If a sequence $X = \{X_n\}$ of fuzzy numbers τ_F – statistically convergent to the fuzzy number X_0, then this sequence statistically converges to the same fuzzy number X_0 with respect to supremum metric d_M.*

Proof. Assume τ_F-st-$\lim X_n = X_0$. Then the set $\{n \in \mathsf{N} \,:\, d_G(X_n, X_0) \not\succ \varepsilon_1\}$ has natural density zero for every $\varepsilon_1 \in \mathsf{F}^+(\mathsf{R})$. Fix $\varepsilon > 0$. Then we have

$$\delta\left(\left\{n \in \mathsf{N} \,:\, \chi_{\left[\left|\underline{X_n}^1 - \underline{X_0}^1\right|,\ \sup_{t\in[\alpha,1]} d\left(X_n^t, X_0^t\right)\right]} \not\succ \varepsilon_1\right\}\right) = 0$$

for every $\alpha \in [0,1]$. It is clear that the inclusion

$$\left\{n \in \mathsf{N} \,:\, \chi_{\left[\left|\underline{X_n}^1 - \underline{X_0}^1\right|,\ \sup_{t\in[\alpha,1]} d\left(X_n^t, X_0^t\right)\right]} \not\succ \varepsilon_1\right\} \supseteq \left\{n \in \mathsf{N} \,:\, \sup_{t\in[\alpha,1]} d\left(X_n^t, X_0^t\right) \ge \varepsilon\right\}$$

holds for every $\alpha \in [0,1]$, i.e., we have

$$\left\{n \in \mathsf{N} \,:\, \chi_{\left[\left|\underline{X_n}^1 - \underline{X_0}^1\right|,\ \sup_{t\in[\alpha,1]} d\left(X_n^t, X_0^t\right)\right]} \not\succ \varepsilon_1\right\} \supseteq \{n \in \mathsf{N} \,:\, d_M(X_n, X_0) \ge \varepsilon\}. \tag{2.1}$$

The inclusion (2.1) say that

$$\delta(\{n \in \mathsf{N} \,:\, d_M(X_n, X_0) \ge \varepsilon\}) = 0$$

ISBN: 978-1-61804-227-9

because left hand side of the inclusion (2.1) has natural density zero.

The converse of this theorem does not valid in general as can be seen by the following example.

Example 2. *Define the sequence* $X=\{X_n\}$ *by*

$$X_n(x):=\begin{cases}\left.\begin{matrix} x-\left(2+\frac{1}{n}\right) & ,\text{if } x\in\left(2+\frac{1}{n},3+\frac{1}{n}\right) \\ \left(4+\frac{1}{n}\right)-x & ,\text{if } x\in[3+\frac{1}{n},4+\frac{1}{n}) \\ 0 & ,\text{otherwise}\end{matrix}\right\} & ,\text{if } n\neq k^2 \ (k\in\mathbb{N}) \\ \left.\begin{matrix} \frac{n(9-x)+1}{n+1} & ,\text{if } x\in[8,9+\frac{1}{n}] \\ 0 & ,\text{otherwise}\end{matrix}\right\} & ,\text{otherwise}\end{cases}$$

and define the fuzzy number X_0 *by*

$$X_0(x):=\begin{cases} x-2 & ,\text{if } x\in(2,3) \\ 4-x & ,\text{if } x\in[3,4). \\ 0 & ,\text{otherwise}\end{cases}$$

Then st-$\lim X_n=X_0$. *But* τ_F *-st-*$\lim X_n$ *does not exist. Now we show this claim. Define*

$$P(x):=\begin{cases} 0 & ,x\in(-\infty,0]\cup[2,\infty) \\ x & ,x\in(0,1] \\ 2-x & ,\text{otherwise}\end{cases}.$$

Then we have

$$\begin{aligned} d_G(X_n,X_0) &= \sup_{\alpha\in[0,1]}\alpha\chi_{\left[\left|\underline{X}_n^1-\underline{X}_0^1\right|,\ \sup_{t\in[\alpha,1]}d\left(X_n^t,X_0^t\right)\right]} \\ &= \sup_{\alpha\in[0,1]}\alpha\chi_{\left[\frac{1}{n},\frac{1}{n}\right]}\end{aligned}$$

if $n\neq k^2$ *(* $k\in\mathbb{N}$ *). Otherwise we have* $d_G(X_n,X_0)=\sup_{\alpha\in[0,1]}\alpha\chi_{\left[0,\frac{1}{n}\right]}$. *Hence we get* $P\nprec d_G(X_n,X_0)$ *if* $n\neq k^2$, *otherwise* $d_G(X_n,X_0)\prec P$. *Consequently we have* τ_F *-st-*$\lim X_n\neq X_0$.

III. τ_F -STATISTICAL LIMIT POINTS

Definition 8 (τ_F -statistical limit point)**.** *Let* $X=\{X_n\}$ *be a sequence of fuzzy numbers and* ν *be a fuzzy number. The number* ν *is called* τ_F *-statistical limit point of the sequence* X *if there is a nonthin subsequence of* X *that* τ_F *-converges to* ν. *We denote the set of all* τ_F *-statistical limit points of the sequence* X *by* $\Lambda_X^{\tau_F}$.

Definition 9 (τ_F -statistical cluster point)**.** *Let* $X=\{X_n\}$ *be a sequence of fuzzy numbers and* μ *be a fuzzy number. The number* μ *is called* τ_F *-statistical cluster point of the sequence* X *if the set*

$$\{n\in\mathbb{N}\ :\ d_G(X_n,\mu)\prec P\}$$

has no natural density zero for every $P\in\mathbf{F}^+(\mathbb{R})$. *Let* $\Gamma_X^{\tau_F}$ *denote the set of all* τ_F *-statistical cluster points of the sequence* X.

Now we give an illustrative example.

Example 3. *Define the sequence* $X=\{X_n\}$ *by*

$$X_n(x):=\begin{cases}\left.\begin{matrix} 1-\frac{nx}{n+1} & ,\text{if } x\in[0,1+\frac{1}{n}] \\ 0 & ,\text{otherwise}\end{matrix}\right\} & ,\text{if } n \text{ is an even number and } n\neq k^2\ (k\in\mathbb{N}) \\ \left.\begin{matrix} \frac{n(9-x)+1}{n+1} & ,\text{if } x\in[8,9+\frac{1}{n}] \\ 0 & ,\text{otherwise}\end{matrix}\right\} & ,\text{if } n \text{ is an even number and } n=k^2\ (k\in\mathbb{N}) \\ \left.\begin{matrix} x-\left(2+\frac{1}{n}\right) & ,\text{if } x\in\left(2+\frac{1}{n},3+\frac{1}{n}\right) \\ \left(4+\frac{1}{n}\right)-x & ,\text{if } x\in[3+\frac{1}{n},4+\frac{1}{n}) \\ 0 & ,\text{otherwise}\end{matrix}\right\} & ,\text{if } n \text{ is an odd number and } n\neq k^2\ (k\in\mathbb{N}) \\ \left.\begin{matrix} x-\left(5+\frac{1}{n}\right) & ,\text{if } x\in[5+\frac{1}{n},6+\frac{1}{n}] \\ 0 & ,\text{otherwise}\end{matrix}\right\} & ,\text{if } n \text{ is an odd number and } n=k^2\ (k\in\mathbb{N})\end{cases}.$$

Define

$$\mu_0(x):=\begin{cases} 1-x & ,\text{if } x\in[0,1] \\ 0 & ,\text{otherwise}\end{cases},$$

$$\nu_0(x):=\begin{cases} x-2 & ,\text{if } x\in(2,3) \\ 4-x & ,\text{if } x\in[3,4), \\ 0 & ,\text{otherwise}\end{cases}$$

$$\gamma_0(x):=\begin{cases} x-5 & ,\text{if } x\in[5,6] \\ 0 & ,\text{otherwise}\end{cases}$$

and

$$\xi_0(x):=\begin{cases} 9-x & ,\text{if } x\in[8,9] \\ 0 & ,\text{otherwise}\end{cases}.$$

Hence we obtain

$$\begin{aligned} L_X &= \{\mu_0,\nu_0,\gamma_0,\xi_0\}, \\ \Lambda_X &= \Gamma_X=\{\mu_0,\nu_0\}, \\ L_X^{\tau_F} &= \{\mu_0,\xi_0\}, \\ \Lambda_X^{\tau_F} &= \Gamma_X^{\tau_F}=\{\mu_0\}.\end{aligned}$$

Since τ_F -convergence implies the convergence with respect to supremum metric d_M, it is clear that $L_X^{\tau_F}\subset L_X$ and $\Lambda_X^{\tau_F}\subset\Lambda_X$. Now we prove the relations between the sets of statistical cluster points:

Theorem 2. *We have* $\Gamma_X^{\tau_F}\subset\Gamma_X$ *for a sequence* $X=\{X_n\}$ *of fuzzy numbers.*

Proof. Take $\mu\in\Gamma_X^{\tau_F}$. By definition, we get

$$\delta(\{n\in\mathbb{N}\ :\ d_G(X_n,\mu)\prec P\})\neq 0$$

for every $P\in\mathbf{F}^+(\mathbb{R})$. Fix $\varepsilon>0$. Then we can write

ISBN: 978-1-61804-227-9

$$\delta\left(\left\{n \in \mathrm{N} : d_G\left(X_n, \mu\right) < \varepsilon_1\right\}\right) \neq 0,$$

since $\varepsilon_1 \in \mathbf{F}^+(\mathrm{R})$. From definition of the metric d_G, we have

$$\delta\left(\left\{n \in \mathrm{N} : \sup_{t \in [\alpha,1]} d\left(X_n^t, \mu^t\right) < \underline{\varepsilon_1}^t = \varepsilon\right\}\right) \neq 0$$

for every $\alpha \in [0,1]$, i.e., we get

$$\delta\left(\left\{n \in \mathrm{N} : d_M\left(X_n, \mu\right) < \varepsilon\right\}\right) \neq 0$$

by the definition of supremum metric d_M. Since the number ε is arbitrary, the proof of theorem is completed.

Theorem 3. *We have* $\Gamma_X^{\tau_F} \subset L_X^{\tau_F}$ *for a sequence* $X = \{X_n\}$ *of fuzzy numbers.*

Proof. Take $\mu \in \Gamma_X^{\tau_F}$. Fix $P \in \mathbf{F}^+(\mathrm{R})$. Then we have $\delta\left(\left\{n(j) \in \mathrm{N} : d_G\left(X_{n(j)}, \mu\right) < P\right\}\right) \neq 0$. Define $\{X\}_K$ by a nonthin subsequence of X such that

$$K = K(P) := \left\{n \in \mathrm{N} : d_G\left(X_n, \mu\right) < P\right\}$$

and $\delta(K) \neq 0$. Then there exists a subset $L \subset K$ such that $\tau_F\text{-}\lim_{n \in L,\, n \to \infty} X_n = \mu$, where the set L has infinitely many elements. Therefore we get $\mu \in L_X^{\tau_F}$.

The converse of this theorem does not hold in general as can be seen in Example 3.

Theorem 4. *We have* $\Lambda_X^{\tau_F} \subset \Gamma_X^{\tau_F}$ *for a sequence* $X = \{X_n\}$ *of fuzzy numbers.*

Proof. Assume $\nu \in \Lambda_X^{\tau_F}$. Then there exists a set $K := \left\{n(j) \in \mathrm{N} : j \in \mathrm{N}\right\}$ such that $\overline{\delta}(K) = l > 0$ and

$\tau_F\text{-}\lim_{j \to \infty} X_{n(j)} = \mu$. Fix $P \in \mathbf{F}^+(\mathrm{R})$. Hence the inclusion

$$\begin{aligned}\left\{n \in \mathrm{N} : d_G\left(X_n, \mu\right) < P\right\} &\supseteq \left\{n(j) \in \mathrm{N} : d_G\left(X_{n(j)}, \mu\right) < P\right\} \\ &= K \setminus \left\{\left\{n(j) \in \mathrm{N} : d_G\left(X_{n(j)}, \mu\right) \geq P\right\} \cup \left\{n(j) \in \mathrm{N} : d_G\left(X_{n(j)}, \mu\right) \nless P\right\}\right\}\end{aligned}$$

holds. Here the sets $\left\{n(j) \in \mathrm{N} : d_G\left(X_{n(j)}, \mu\right) \geq P\right\}$ and $\left\{n(j) \in \mathrm{N} : d_G\left(X_{n(j)}, \mu\right) \nless P\right\}$ has finite many elements. Hence we have

$$\begin{aligned}\overline{\delta}\left(\left\{n \in \mathrm{N} : d_G\left(X_n, \mu\right) < P\right\}\right) &\geq \overline{\delta}(K) \\ &- \overline{\delta}\left(\left\{n(j) \in \mathrm{N} : d_G\left(X_{n(j)}, \mu\right) \geq P\right\} \cup \left\{n(j) \in \mathrm{N} : d_G\left(X_{n(j)}, \mu\right) \nless P\right\}\right) \\ &= l.\end{aligned}$$

Therefore we get $\delta\left(\left\{n \in \mathrm{N} : d_G\left(X_n, \mu\right) < P\right\}\right) \neq 0$.

REFERENCES

[1] Y. Altın, M. Et, and R. Çolak, "Lacunary statistical and lacunary strongly convergence of generalized difference sequences of fuzzy numbers", *Comput. Math. Appl.*, vol. 52, pp. 1011-1020, 2006.

[2] H. Altınok, R. Çolak, and M. Et, "λ -difference sequence spaces of fuzzy numbers", *Fuzzy Sets and Systems*, vol. 160, no. 21, pp. 3128-3139, 2009.

[3] H. Altınok, Y. Altın, and M. Işık, "Statistical convergence and strong p-Cesàro summability of order β in sequences of fuzzy numbers, *Iran. J. Fuzzy Syst.*, vol. 9, no. 2, pp. 63-73, 2012.

[4] S. Aytar, "Statistical limit points of sequences of fuzzy numbers", *Information Sciences*, vol. 165, pp. 129-138, 2004.

[5] S. Aytar, "A neighbourhood systems of fuzzy numbers and its topology", submitted.

[6] P. Diamond, and P. Kloeden, "Metric Spaces of Fuzzy Sets: Theory and Applications", World Scientific, Singapore, 1994.

[7] D. Dubois, and H. Prade, "Operations on fuzzy numbers", *Int. J. Systems Science*, vol. 9, pp. 613-626, 1978.

[8] J-x. Fang, and H. Huang, "On the level convergence of a sequence of fuzzy numbers", *Fuzzy Sets and Systems*, vol. 147, pp. 417-435, 2004.

[9] Z. Guangquan, "Fuzzy distance and fuzzy limit of fuzzy numbers", *Busefal*, vol. 33, pp. 19-30, 1987.

[10] Z. Guangquan, "Fuzzy continuous function and its properties", *Fuzzy Sets and Systems*, vol. 43, pp. 159-171, 1991.

[11] J. Hančl, L. Mišík and J. T. Tóth, "Cluster points of sequences of fuzzy real numbers", *Soft Computing*, vol. 14, no. 4, pp. 399-404, 2010.

[12] O. Kaleva, "On the convergence of fuzzy sets", *Fuzzy Sets and Systems*, vol. 17, pp. 53-65, 1985.

[13] M. Matloka, "Sequences of fuzzy numbers", *Busefal*, vol. 28, pp. 28-37, 1986.

[14] M. Mizumoto, and K. Tanaka, "The four operations of arithmetic on fuzzy numbers", *Systems-Computers-Controls*, vol. 7, pp. 73-81, 1976.

[15] M. Mizumoto, and K. Tanaka, "Some properties of fuzzy numbers", *Advances in fuzzy set theory and applications*, pp. 153-164, North-Holland, Amsterdam-New York, 1979.

[16] S. Nanda, "On sequence of fuzzy numbers", *Fuzzy Sets and Systems*, vol. 33, pp. 123-126, 1989.

[17] F. Nuray, and E. Şavaş, "Statistical convergence of sequences of fuzzy numbers", *Math. Slovaca*, vol. 45, no. 3, pp. 269-273, 1995.

[18] M.L. Puri, and D.A. Ralescu, "Fuzzy random variables", *J. Math. Anal. Appl.*, vol. 114, pp. 409-422, 1986.

[19] Ö . Talo, and F. Başar, "Determination of the duals of classical sets of sequences of fuzzy numbers and related matrix transformations", *Comput. Math. Appl.* Vol. 59, pp. 717-733, 2009.

[20] H. Steinhaus, "Sur la convergence ordinaire et la convergence asymptotique", *Colloq. Math.*, vol. 2, pp. 73-74, 1951.

Ulaş¸ Yamancı received the MSc at Graduate School of Natural and Applied Sciences at Süleyman Demirel University of Isparta in Turkey. His research interests are: Toeplitz operator, Berezin symbols, Reproducing Kernels, Statistical convergence, Ideal convergence.

M. Gürdal received the PhD degree in Mathematics for Graduate School of Natural and Applied Sciences at Süleyman Demirel University of Isparta in Turkey. His research interests are in the areas of functional analysis and operator theory including statistical convergence, Berezin sembols, Banach algebras, Toeplitz Operators. He has published research articles in reputed international journals of mathematical science. He is referee and editor of mathematical journals.

S. Aytar received the PhD degree in Mathematics for Graduate School of Natural and Applied Sciences at Süleyman Demirel University of Isparta in Turkey. His research interests are in the areas of functional analysis, convex analysis, applied mathematics including statistical convergence, fuzzy sets, rough convergence. He has published research articles in reputed international journals of mathematical science. He is a referee of mathematical journals.

One approach to solve some problems of management under uncertainty

Teimuraz Tsabadze
a) Georgian Technical University
b) Georgian American University
Tbilisi, Georgia
Teimuraz.Tsabadze@yahoo.com

Archil Prangishvili
Rector
Georgian Technical University
Tbilisi, Georgia
rectoroffice@gtu.ge

Abstract— We consider the approaches to making decisions for control problems in nonstandard situations with a lack of the previous experience and incomplete knowledge of the considered problem. In such cases we usually cannot do without expert evaluations which lead to the process of group decision-making, and it becomes necessary to solve a problem of alternatives aggregation. It has been proposed to solve such problems by means of fuzzy sets. The approach is based on the coordination index and the similarity of finite collections of fuzzy sets and takes into account the specific character of the fuzzy aggregation operator. The approach is discussed and its algorithm is presented. An example of the application of the proposed method is given.

Keywords— Group decision-making; finite collection of fuzzy set;, increasing shuffling; coordination index; similarity; fuzzy aggregation operator.

I. INTRODUCTION

The modern system of knowledge about management (control) is formed on the basis of various sciences. The management process is a difficult complex system. for example in management of marketing. An important task of management is the integration of all qualities and aspects of the controlled object in order to achieve general purposes set before this object [6] – this means the necessity to analyze a multifactor function whose parameter values are frequently represented by fuzzy or incomplete information which is as a rule is subjective.

The control process is entirely based on the decision-making factor.

According to the character of problems, which arise in the course of activity of a controlled object, and methods used for decision-making, we can distinguish the groups of problems [3,4] as shown in the table below:

Structured or, following the terminology of [5], programmed problems - These are solutions which once made become the rules controlling all future actions. They are called structured (programmed) decisions and are part of daily activities of a controlled object, are permanently repeated and comply with the policy adopted by managers. They can be obtained by economical-mathematical methods and the algorithms, already developed, are available in most cases.

Non-structured (non-programmed) problems - A solution should be obtained in a non-standard situation with a lack of previous experience is characterized by incomplete knowledge about the main components and needs non-ordinary approaches. In this area heuristic methods are of great use. Such solutions cause embarrassment in the manager because each time he has to seek for a procedure of their selection. This group includes in particular solutions for a concrete controlled object to get out of crisis, as well as solutions concerning a search for a strategy, determination of an amount of capital to be invested into new production, and so on.

Scientifically substantiated standard problems solved by the statutory rules on solution selection - Scientifically substantiated solutions are those which are adopted on the basis of scientific arguments. They are also called rational because they are supported by scientifically proven schemes incorporating various models, comparisons of different variants and so on. Obtaining such a solution needs much time, the process is frequently iterative, but enables the manager to find a solution.

Intuitive problems and problems based on opinions - As different from scientifically substantiated solutions, intuitive solutions are taken promptly, by feeling that one must act this way and not differently. Such solutions are based on personal experience.

It is obvious that managers encounter big difficulties when they are confronted with non-structured problems (item 2 in the above table), which frequently have to be solved by heuristic methods.

In the present paper, we focus the attention on “the weakest link” – the non-structured process of making management decisions. Our objective is to solve the most frequently occurring problems in situations which lack the previous experience and original information is incomplete. In our opinion confirmed by experience, after the detailed analysis of the management situation the next important step is to find input data to be used in making a final decision. Here, as a rule, we cannot do without experts’ evaluations which lead to the process of group decision-making. In this situation, the manager (managers) has (have) to solve the problem of alternatives aggregation.

All people usually think in uncertain categories and are not guided only by the "yes" or "no" principle. For example, if the question is asked "Is this project profitable?", classical mathematics answers either "yes" or "no", while the relatively young fuzzy sets theory gives one a chance to model such categories as "unprofitable", "not quite profitable", "profitable", "more profitable", "very profitable" and so on.

In 1965, American scientist L. Zadeh published the paper "Fuzzy Sets" [10], in which he drew the attention of the world scientific community to an absolutely novel direction in the development of mathematics and applied sciences. Instead of the classical membership function of some object in some set (0 or 1), Zadeh introduced the continuous membership interval – [0; 1].

Fuzzy sets theory made an invaluable contribution to the development of new information-control technologies. It is a powerful instrument of analysis of weakly structured, fragmentary, incomplete and fuzzy information.

The point is that in the absence of a general collection of data, i.e. of a sufficiently large initial database, even such well-tested instruments of modeling various situations as *probability theory* and *mathematical statistics* are not capable to fairly take into consideration incomplete and fuzzy information [2].

We propose an approach for group-decision making under uncertainty, where experts' opinions are expressed by quantitative values.

II. Essential notions and the theoretical background

Let X be the universe. A mapping $A: X \to [0;1]$ is called *a fuzzy set* on X. The value $\mu_A(x)$ of A at $x \in X$ denotes the membership degree of x in A. Thus the fuzzy set A can be determined as follows:

$$A = \{(x \mid \mu_A(x))\}, \ x \in X, \ \mu_A(x) \in [0;1].$$

$\Psi(X) = \{\mu \mid \mu : X \to [0;1]\}$ is the lattice of all fuzzy sets on X.

$\varnothing$ is a minimal element of $\Psi(X)$: $\mu_\varnothing(x) = 0 \quad \forall x \in X$.

U is a maximal element of $\Psi(X)$: $\mu_U(x) = 1 \quad \forall x \in X$.

$$A = B \Leftrightarrow \mu_A(x) = \mu_B(x) \ \forall x \in X, \quad A, B \in \Psi(X).$$

$$A \subseteq B \Leftrightarrow \mu_A(x) \le \mu_B(x) \ \forall x \in X, \quad A, B \in \Psi(X).$$

$A \subset B \Leftrightarrow \mu_A(x) \le \mu_B(x) \quad \forall x \in X$ and

$$\exists x_0 \in X \mid \mu_A(x_0) < \mu_B(x_0)$$

The union of fuzzy sets A and B is $\mu_{A \cup B}(x) = \max\{\mu_A(x), \mu_B(x)\} \ \forall x \in X$.

The intersection of fuzzy sets A and B is $\mu_{A \cap B}(x) = \min\{\mu_A(x), \mu_B(x)\} \quad \forall x \in X$.

It is not difficult to verify that the distributivity of $\cap$ and $\cup$ holds in $\Psi(X)$:

$$\begin{aligned} A \cap (B \cup C) &= (A \cap B) \cup (A \cap C), \\ A \cup (B \cap C) &= (A \cup B) \cap (A \cup C). \end{aligned} \tag{2.1}$$

We say that the function $v\colon \Psi(X) \to \Re^+$ is *the isotone valuation on* $\Psi(X)$ if [1]:

$$v(A \cup B) + v(A \cap B) = v(A) + v(B)$$

and (2.2)

$$A \subseteq B \Rightarrow v(A) \le v(B).$$

We say that the isotone valuation v is *continuous* if for each $a \in [v(\varnothing); v(U)]$ there exists

$A \in \Psi(X)$ such that $v(A) = a$.

Let us consider the equation

$$\rho(A,B) = v(A \cup B) - v(A \cap B). \tag{2.3}$$

We will show that (2.3) represents the *metric* on $\Psi(X)$, i.e. it meets the following requirements:

1) $\rho(A,B) = 0 \Leftrightarrow A = B$;

2) $\rho(A,B) = \rho(B,A)$;

3) $\rho(A,C) + \rho(C,B) \ge \rho(A,B), \quad \forall C \in \Psi(X)$.

Proof. Let $\rho(A,B) = 0 \Rightarrow v(A \cup B) = v(A \cap B)$. By (2.4) we can easily conclude that $A \cup B = A \cap B \Rightarrow A = B$.

Now let $A = B \Rightarrow A \cup B = A \cap B \Rightarrow v(A \cup B)$

$= v(A \cap B) \Rightarrow \rho(A,B) = 0$

So, 1) is true.

$$\rho(A,B) = v(A \cup B) - v(A \cap B)$$
$$= v(B \cup A) - v(B \cap A) = \rho(B,A)$$

2) is also true.

$$\rho(A,B) = v(A \cup B) - v(A \cap B)$$

$$\overset{(2.2)}{\le} v((A \cup C) \cup (C \cup B)) - v((A \cap C) \cap (C \cap B))$$

$$\overset{(2.2)}{=} v(A \cup C) + v(C \cup B) - v((A \cup C) \cap (C \cup B))$$

$$- v(A \cap C) - v(C \cap B) + v((A \cap C) \cup (C \cap B))$$

$$\overset{(2.1),(2.2)}{=} \rho(A,C) + \rho(C,B) + v(C \cap (A \cup B))$$

$$- v(C \cup (A \cap B)) \overset{(2.2)}{\le} \rho(A,C) + \rho(C,B)$$

ISBN: 978-1-61804-227-9

Thus 3) has been proved. □

Ψ (X) with isotone valuation v and metric (2.3) is called *the metric lattice* of fuzzy sets.

Definition 2.1[7]. In the metric lattice the fuzzy set A^* is the representative of a finite collection of *fuzzy* sets $\{ A_j \}$, $j = \overline{1,m}$, m=2,3,…, if

$$\sum_{j=1}^{m} \rho(A^*, A_j) \le \sum_{j=1}^{m} \rho(B, A_j), \quad \forall B \in \Psi(X). \tag{2.4}$$

Now let us introduce the concept of increasing shuffling of a finite collection of fuzzy sets.

Definition 2.2[7]. A finite collection of fuzzy sets $\{ A_j^{'} \}$ is the increasing shuffling of a finite collection of fuzzy sets $\{ A_j \}$ if for each $x \in X$ the finite sets $\{ \mu_{A_j}(x) \}$ and $\{ \mu_{A_j^{'}}(x) \}$ are equal to each other and $\mu_{A_1^{'}}(x) \le \mu_{A_2^{'}}(x) \le \ldots \le \mu_{A_m^{'}}(x)$, $j = \overline{1,m}$, $m = 2,3,\ldots$.

Thus the increasing shuffling represents a finite collection of *nested* fuzzy sets. The equality

$$\sum_{j=1}^{m} \rho(B, A_j) = \sum_{j=1}^{m} \rho(B, A_j^{'}) \tag{2.5}$$

holds in the metric lattice for any $B \in \Psi$ (X) and a finite collection of fuzzy sets *{* A_j *}*, $j = \overline{1,m}$, $m = 2,3,\ldots$.

Theorem 2.1[7]. *In the metric lattice of fuzzy sets the representative* A^* *of a finite collection of fuzzy sets* $\{ A_j \}$, $j = \overline{1,m}$, m=2,3,…, *is determined as follows:*

$$A_{m/2}^{'} \subseteq A^* \subseteq A_{m/2+1}^{'} \quad \textit{if m is even;}$$

$$A^* = A_{(m+1)/2}^{'} \quad \textit{if m is odd}$$

Here and further the symbol [ ] denotes the integer part of a number.

We say that a finite collection of fuzzy sets is *symmetrical* [7] if in its increasing shuffling the first [(2m+1)/4] sets are equal to $\varnothing$ and the last [(2m+1)/4] sets are equal to U, m=2,3,… .

Definition 2.3[7]. In a metric lattice of fuzzy sets the functional

$$S: \underbrace{\Psi(X) \times \ldots \times \Psi(X)}_{m \ times} \to \mathbb{R}^+$$

is the coordination index of the finite collection of fuzzy sets $\{ A_j \}$, $j = \overline{1,m}$, m=2,3,…, if it satisfies the following postulates*:*

P1. $S\{A_j\} = 0$ if and only if the finite collection of fuzzy sets $\{ A_j \}$ is symmetrical;

P2. $S\{A_j\}$ reaches a maximal value if and only if all fuzzy sets of the finite collection are equal to one another;

P3. $S\{A_j\} \ge S\{B_j\}$ if $\sum_{j=1}^{m} \rho(A^*, A_j) \le \sum_{j=1}^{m} \rho(B^*, B_j)$; moreover, $S\{A_j\} = S\{B_j\}$ if and only if $\sum_{j=1}^{m} \rho(A^*, A_j) = \sum_{j=1}^{m} \rho(B^*, B_j)$;

P4. $S\{A_j \cup B_j\} + S\{A_j \cap B_j\} = S\{A_j\} + S\{B_j\}$.

Setting aside the question of independence and completeness of the postulates introduced above (Definition 2.3), by constructing the corresponding example [7] we show that this axiomatic is consistent.

Theorem 2.2[7]. *In the metric lattice of fuzzy sets the functional* $S\{A_j\}$ *is the coordination index of the finite collection of fuzzy sets* $\{ A_j \}$, $j = \overline{1,m}$, m=2,3,…, *if*

$$S\{A_j\} = q(\rho(\varnothing, U) - [(2m+1)/4]^{-1} \times \sum_{j=1}^{m} \rho(A^*, A_j)), q > 0 \tag{2.6}$$

Moreover, if the isotone valuation v is continuous, this representation is unique.

It is obvious that

$$S_{\max} = q\rho(\varnothing, U), \quad \theta > 0. \tag{2.7}$$

A method of fuzzy aggregation based on group expert evaluations is presented in [8]. Below we give the basic theoretical results and explanations needed to understand the principles and work of this method. First, we introduce the important concept of similarity for finite collections of fuzzy sets, where the metric approach is used.

Definition 2.4[8]. In the metric lattice of fuzzy sets the finite collection of fuzzy sets $\{A_j\}$ is similar to the finite collection of fuzzy sets $\{B_j\}$ if for each $x \in X$ $\rho(A_i^{'}, A_{i-1}^{'}) = k\rho(B_i^{'}, B_{i-1}^{'})$, $i = \overline{2,m}$, $j = \overline{1,m}$, $m = 2,3,\ldots$, where $k > 0$ is the similarity coefficient, $\{A_j^{'}\}$, $\{B_j^{'}\}$ are the increasing shufflings of $\{A_j\}$ and $\{B_j\}$ respectively.

We denote the similarity of two finite collections of fuzzy sets in the metric lattice of fuzzy sets by $\{A_j\} \overset{k}{\cong} \{B_j\} \Leftrightarrow \{B_j\} \overset{1/k}{\cong} \{A_j\}$ or simply by $\{A_j\} \cong \{B_j\}$.

One way of constructing such a finite collection of fuzzy sets in the metric lattice with continuous isotone valuation v is presented in [8]. Let $\{B_j\}$, $j = \overline{1,m}$, m=2,3,…, be a finite collection of fuzzy sets. To construct a finite collection of fuzzy sets $\{C_j\}$ similar to $\{B_j\}$, the following general formula are used:

$$\begin{aligned} &\forall x \in X \quad \mu_{C_j}(x) = k\rho(B_1^{'}, B_j^{'}) + \mu_{C_1}(x), \\ &j = \overline{1,m},\ m = 2,3,\ldots,\ \mu_{C_1}(x) = v(C_1) = \alpha \mid 0 \le \alpha < 1 \end{aligned} \quad . \tag{2.8}$$

Now we present an important theorem. We remind that the maximal coordination index $S_{\max}$ of a finite collection of fuzzy sets is determined by (2.7).

Theorem 2.3 [8]. *If* $\{A_j\}\overset{k}{\cong}\{B_j\}$, *then for the coordination indices of these two finite collections of fuzzy sets the equality* $S\{A_j\}=kS\{B_j\}+(1-k)S_{\max}$, $j=\overline{1,m}$ $m=2,3,...$ *holds.*

Corollary. $\{A_j\}\overset{1}{\cong}\{B_j\}\Rightarrow S\{A_j\}=S\{B_j\}$, $j=\overline{1,m}$, $m=2,3....$

The following theorem is the main theoretical basis of our fuzzy aggregation method.

Theorem 2.4 [8]. *In the metric lattice of fuzzy sets with continuous isotone valuation for any two finite collection of fuzzy sets* $\{A_j\}$ *and* $\{B_j\}$ *such that* $S\{A_j\}$, $S\{B_j\}$ < $S_{\max}$ *there exists a finite collection of fuzzy sets* $\{C_j\}$ *such that* $\{C_j\}\cong\{B_j\}$ *and* $S\{C_j\}=S\{A_j\}$, $j=\overline{1,m}$, m=2,3,… .

The next theorem determines some specific conditions for uniqueness of finite collection of fuzzy sets, which is similar to the given one.

Theorem 2.5 [8]. *In the metric lattice of fuzzy sets with continuous isotone valuation for any two finite collections of fuzzy sets such as* $\{C_j\}\overset{k}{\cong}\{B_j\}$ *and* $S\{C_j\} > S\{B_j\}$, *for each k there exists a unique finite collection of fuzzy sets* $\{A_j\}$ *such that* $\{A_j\}\overset{1}{\cong}\{C_j\}$ *and* $A_l=B_l^{'}$, $j=\overline{1,m}$, $l\in\{1,2,...,m\}$, $m=2,3...$.

By Theorem 2.5 we obtain the general formula

$$\mu_{A_{lj}}(x)=\mu_{B_l^{'}}(x)+k(v(B_j^{'})-v(B_l^{'})),\quad j=\overline{1,m},\ l\in\{1,2,...,m\},\ \forall x\in X \tag{2.9}$$

In the sequel, we will need one specific modification of (2.9). Under the conditions of Theorem 2.5 let us consider the finite collection of fuzzy sets $\{\bar{A}_j\}$ whose membership functions are the arithmetic means of the respective membership functions (2.9):

$$\{\bar{A}_j\}=\left\{\frac{\sum_{l=1}^{m}\mu_{A_{lj}}(x)}{m}\right\},\ j=\overline{1,m},\ m=2,3,...,\ \forall x\in X. \tag{2.10}$$

From the last expression, by (2.9) $\{\bar{A}_j\}=\left\{\frac{\sum_{l=1}^{m}\left(\mu_{B_l^{'}}(x)-kv\left(B_l^{'}\right)\right)}{m}+kv\left(B_j^{'}\right)\right\}$ and using the notation

$$c=\frac{\sum_{l=1}^{m}\left(\mu_{B_l^{'}}(x)-kv\left(B_l^{'}\right)\right)}{m}, \tag{2.11}$$

we obtain that

$$\{\bar{A}_j\}=\left\{c+kv\left(B_j^{'}\right)\right\},\ j=\overline{1,m}, l\in\{1,2,...,m\}, m=2,3,...,\forall x\in X \tag{2.12}$$

It is obvious that for $\forall x\in X$ $\{\bar{A}_j\}$ is a finite collection of nested fuzzy sets. From (2.1), (2.10), (2.12), Definition 2.4 and Theorem 2.5 it follows that $\rho(\bar{A}_i,\bar{A}_{i-1})=v(\bar{A}_i)-v(\bar{A}_{i-1})=c+kv(B_i^{'})-c-kv(B_{i-1}^{'})$ $=k\rho(B_i^{'},B_{i-1}^{'})=\rho(C_i,C_{i-1})\Rightarrow\{\bar{A}_j\}\overset{1}{\cong}\{C_j\}$, $i=\overline{2,m}$, $j=\overline{1,m}$, $\forall x\in X$. According to the corollary of Theorem 2.3, $S\{\bar{A}_j\}=S\{C_j\}$.

Thus we have proved

Proposition 2.1[8]. *If under the conditions of Theorem* 2.5 *the finite collection of fuzzy sets* $\{\bar{A}_j\}$ *is determined by* (2.10), *then* $\{\bar{A}_j\}\overset{1}{\cong}\{C_j\}$ *and, consequently,* $S\{\bar{A}_j\}=S\{C_j\}$, $j=\overline{1,m}$, $m=2,3,...,$ $\forall x\in X$.

For the realization of the proposed method we need a specific aggregation operator that meets certain requirements. Presently, in the fuzzy set theory there are several well known fuzzy aggregation operators (see e.g. [9]). The representative of a finite collection of fuzzy sets determined by Definition 2.1 is a new kind of fuzzy aggregation operator. By Theorem 2.1 the representative of a finite collection of fuzzy sets $\{A_j\}$, $j=\overline{1,m}$, m=2,3,…, yields the expression $A_{m/2}^{'}\subseteq A^*\subseteq A_{m/2+1}^{'}$ when m is even or $A^*=A_{(m+1)/2}^{'}$ when m is odd. This means that if m is even and $A_{m/2}^{'}\subset A_{m/2+1}^{'}$, then the representative can take an infinite number of values. In [8], after the detailed discussion and justification, the following form of a unique representative is defined (here and in the sequel we use the notation μ_A instead of $\mu_A(x)$, $\forall x\in X$):

$$\mu_{A^*}=\begin{cases}(\mu_{A_{[m/2]}^{'}}+\mu_{A_{[(m+3)/2]}^{'}})/2 \ \ \textit{if} \sum\limits_{j=1}^{[(m+1)/2]}\rho(A_j^{'},A_{[m/2]}^{'})=\sum\limits_{j=[m/2]+1}^{m}\rho(A_j^{'},A_{[(m+3)/2]}^{'}),\\ \mu_{A_{[m/2]}^{'}}+\dfrac{\sum\limits_{j=1}^{[(m+1)/2]}\rho(A_j^{'},A_{[m/2]}^{'})}{\sum\limits_{j=1}^{[(m+1)/2]}\rho(A_j^{'},A_{[m/2]}^{'})+\sum\limits_{j=[m/2]+1}^{m}\rho(A_j^{'},A_{[(m+3)/2]}^{'})}(\mu_{A_{[(m+3)/2]}^{'}}-\mu_{A_{[m/2]}^{'}}) \ \ \textit{otherwise.}\end{cases} \tag{2.13}$$

We finish our theoretical background by determining one specific modification of the introduced fuzzy aggregation operator, which is obtained from the interrelation between the sums in (2.13) and the condition of similarity of two finite collections of fuzzy sets.

ISBN: 978-1-61804-227-9

Lemma 2.1 [8] *If* $\{A_j\} \overset{k}{\cong} \{B_j\}$, $j = \overline{1,m}$, *m=2,3,… then*

$$\sum_{j=1}^{[(m+1)/2]} \rho(A_j^{'}, A_{[m/2]}^{'}) = k \sum_{j=1}^{[(m+1)/2]} \rho(B_j^{'}, B_{[m/2]}^{'}) \qquad and$$

$$\sum_{j=[m/2]+1}^{m} \rho(A_j^{'}, A_{[(m+3)/2]}^{'}) = k \sum_{j=[m/2]+1}^{m} \rho(B_j^{'}, B_{[(m+3)/2]}^{'}) .$$

Assume that the finite collection of fuzzy sets $\{\overline{A}_j\}$ is given by (2.10). We want to determine $\mu_{\overline{A}_{[m/2]}}$ and $\mu_{\overline{A}_{[(m+3)/2]}}$. By (2.12) and (2.11) we easily obtain that

$$\mu_{\overline{A}_{[m/2]}} = c + kv(B_{[m/2]}^{'}); \mu_{\overline{A}_{[(m+3)/2]}} = c + kv(B_{[(m+3)/2]}^{'}) \qquad (2.14)$$

Using (2.3) and (2.14) it is easy to check that the equality

$$\mu_{\overline{A}_{[(m+3)/2]}} - \mu_{\overline{A}_{[m/2]}} = k\rho(B_{[m/2]}^{'}, B_{[(m+3)/2]}^{'}) \qquad (2.15)$$

holds.

Finally, if the conditions of Theorem 2.5 are satisfied, $\{\overline{A}_j\}$ is given by (2.10) and $\{\overline{A}_j\} \overset{k}{\cong} \{B_j^{'}\}$, then by Lemma 2.1, (2.14) and (2.15) we can rewrite formula (2.13) as

$$\mu_{\overline{A}_j} = \begin{cases} c + k\dfrac{v(B_{[m/2]}^{'}) + v(B_{[(m+3)/2]}^{'})}{2} \quad if \quad \sum\limits_{j=1}^{[(m+1)/2]} \rho(B_j^{'}, B_{[m/2]}^{'}) = \sum\limits_{j=[m/2]+1}^{m} \rho(B_j^{'}, B_{[(m+3)/2]}^{'}), \\ c + k(v(B_{[m/2]}^{'}) + \dfrac{\rho(B_{[m/2]}^{'}, B_{[(m+3)/2]}^{'}) \sum\limits_{j=1}^{[(m+1)/2]} \rho(B_j^{'}, B_{[m/2]}^{'})}{\sum\limits_{j=1}^{[(m+1)/2]} \rho(B_j^{'}, B_{[m/2]}^{'}) + \sum\limits_{j=[m/2]+1}^{m} \rho(B_j^{'}, B_{[(m+3)/2]}^{'})}) \quad otherwise. \end{cases} \qquad (2.16)$$

III. DESCRIPTION OF THE FUZZY AGGREGATION METHOD AND ITS ALGORITHM

It is assumed here that all experts have the same qualification (otherwise see Remark 3.1). Let the group of experts evaluate the membership degree of the fuzzy object in the given universe. As a result of this evaluation we obtain a finite collection of fuzzy sets and proceed to the process of group decision-making. We try to construct the fuzzy set of the obtained finite collection of fuzzy sets. Assume that we can calculate the coordination index of the obtained finite collection of fuzzy sets at each element of the universe. It is clear that the experts' evaluations at each element form a finite collection of one-element fuzzy sets. If there existed several elements, where the coordination index reaches its maximal value, then by (2.6) all members of the finite collection of one-element fuzzy sets would be equal to one another. This means that then the experts' work is ideal and each of the fuzzy sets constructed by them is the only result of group decision-making. But by the nature of fuzziness such a theoretical chance may occur very rarely.

We are interested in the cases where the coordination index does not reach its maximal value. Suppose that the greatest value of the coordination index is reached at one element of the universe (there may also be a few elements). In that case we consider it to be the experts' best attempt! They are not automata and cannot act with invariably constant concentration on each element of the universe. Let us take an analogy from sport. In such kinds of sports as long and high jumps, discus throwing, weight lifting and so on a sportsman is evaluated by his best result in a finite series of attempts. In our case we consider that this element is the *point of maximal coordination* of the experts and suppose that they can potentially act with the same coordination at all other elements of the universe. We obtain the result of fuzzy aggregation at the point of maximal coordination by using operator (2.13). Let us try to map this best attempt onto the other elements of the universe.

The way of implementing this idea is as follows. We choose an element of the universe where the value of the coordination index is smaller than the greatest value. If the experts' evaluations at this element are not a finite collection of nested one-element fuzzy sets, then we operate with its increasing shuffling. Thus, without loss of generality this finite collection can be assumed to consist of nested one-element fuzzy sets. Let us construct such a finite collection of one-element fuzzy sets that is "similar" to the above-mentioned finite collection of nested one-element sets and at the same time has the greatest coordination index. Now let us discuss the notion of "similarity" in informal terms.

We can imagine a finite collection of one-element fuzzy sets as a finite set of geometric points with equal abscissae, while their ordinates as representing the evaluations of experts at this element of the universe. Then any finite collection of one-element fuzzy sets similar to the above-mentioned one can be represented by the same number of points with equal abscissae, while the distances between the neighboring ordinates are proportional to the respective distances of the finite collection of one-element fuzzy sets. In the other words, in a similar finite collection of one-element fuzzy sets the ordinates are close to or distant from one another, preserving at the same time the proportionality of the respective distances. The strict mathematical determination of the similarity for two finite collections of fuzzy sets is given in Definition 2.4. Formula (2.8) allows us to construct a finite collection of one-element fuzzy sets similar to the finite collection of nested one-element sets.

From Definition 2.4 and formula (2.8) it follows that there exists an infinite set of finite collections of one-element fuzzy sets similar to the given one. As mentioned above, we have to choose a finite collection of one-element fuzzy sets, the coordination index of which is equal to the coordination index of the finite collection of one-element fuzzy sets constructed at the point of maximal coordination. By Theorem 2.4 we can construct a finite collection of one-element fuzzy sets which is similar to the given one and, at the same time, has the greatest coordination index. But even in that case there exists an infinite set of such finite collections of one-element fuzzy sets. Here the problem reduces to determining the unique finite collection of one-element fuzzy sets which takes into proper account the opinions of all experts *pari passu*.

ISBN: 978-1-61804-227-9

Not to "offend" anyone of the experts we act the following way: choose from this infinite set such a finite collection of one-element fuzzy sets where the first one-element fuzzy set is equal to the first one-element fuzzy set of the given finite collection of one-element fuzzy sets; next choose such a finite collection of one-element fuzzy sets where the second one-element fuzzy set is equal to the second one-element fuzzy set of the given finite collection of one-element fuzzy sets and so on; finally, choose such a finite collection of one-element fuzzy sets where the last one-element fuzzy set is equal to the last one-element fuzzy set of the given finite collection of one-element fuzzy sets. It is obvious that the number of constructed finite collections of one-element fuzzy sets is equal to the number of experts.

In the sequel we will apply the mapping $\mu : X \to [0;b] \subset \mathbb{R}$ (here we use the membership function with interval [0;b] instead of the standard interval standard [0;1] in order to make the examples more obvious).

Remark 3.1. The method can also be applied in the case of different qualification of experts (for example by using weight coefficients). The essence and structure of the proposed approach will not be affected.

Let us summarize briefly the stages of the proposed method and give its formal algorithm.

Let $\Psi(X) = \{\mu \mid \mu : X \to [0;b] \subset \Re\}$ be the metric lattice with continuous isotone valuation *v*, the universe *X* be the finite set $\{x_1, x_2, ..., x_N\}$ and the group of *m* experts evaluate the degree of membership of the concept *B* in the universe. As a result we have the finite collection of fuzzy sets: $\{ B_j \}$, $j = \overline{1,m}$, m=2,3,… .

First we construct its increasing shuffling $\{B_j^{'}\}$ and under assumption *q* = 1 compute $S_{\max}$ by (2.5) . After that we compute the coordination index value at each element of the universe by (2.4). Here we have several alternatives.

Assume that at some element the value of this index is maximal (there can be several such elements)**.** This means that each of *m* constructed fuzzy sets is the result of fuzzy aggregation for these elements.

Now consider only such values of coordination indices that are smaller than the greatest value. Choose among them such an element, where the coordination index value is the greatest (there can be several such elements). It is the point of maximal coordination. Denote by $\{A_j\}$ the finite collection of one-element fuzzy sets at this element. In this case the result of fuzzy aggregation is obtained by (2.13).

Now consider such elements where the coordination index values are smaller than the greatest value. Let demonstrate the operation of the algorithm at one of these elements. Suppose that at an element $x \in X$ we have the finite collection of one-element fuzzy sets $\{B_j^{'}\}$. Then by Theorem 2.4 we construct at this element a finite collection of one-element fuzzy sets $\{C_j\}$ such that $\{C_j\} \cong \{B_j^{'}\}$ and $S\{C_j\}=S\{A_j\}$, $j = \overline{1,m}$, m=2,3,… . We know that there exists an infinite set of such finite collections of one-element fuzzy sets. Theorem 2.5 allows us to choose from this set exactly such *m* finite collections of one-element fuzzy sets where each of them has common elements with $\{B_j^{'}\}$. Taking into account the opinion of each expert, we take as a result such a finite collection of one-element fuzzy sets where each member is the arithmetic mean of respective members of these *m* unique finite collections of one-element fuzzy sets (we remind that each member is a real number). To obtain this finite collection of one-element fuzzy sets we use formula (2.10). By Proposition 2.1 this finite collection of one-element fuzzy sets is similar to $\{C_j\}$ with the similarity coefficient $k = 1$ and, by Corollary of Theorem 2.3, has the greatest coordination index.

Thus the constructed finite collection of one-element fuzzy sets satisfies the conditions of Proposition 2.1 and we obtain the result of fuzzy aggregation by operator (2.16). In the same way one can obtain a fuzzy aggregation operator at any other element of the given universe.

Remark 3.2. We do not need to carry out long and tedious calculations for constructing *m* new finite collections of one-element fuzzy sets with each of them having common elements with $\{B_j^{'}\}$ and determining their arithmetic means. As a matter of fact, via the coefficient of similarity we operate only with the given finite collection of one-element fuzzy sets $\{B_j^{'}\}$. In our opinion it is an effective and nice circumstance.

Now we present the generalized algorithm of the fuzzy aggregation method.

Algorithm

Step 0: *Initialization: the finite collection of one-element fuzzy sets* $\{B_j\}$, *its increasing shuffling* $\{B_j^{'}\}$, $j = \overline{1,m}$, $m = 2,3,...$. *Denote the result of the fuzzy aggregation in element* x_i, $i = \overline{1,N}$ *by* $\mu(x_i)$.

Step 1: *Compute the values of coordination indices of the finite collection of one-element fuzzy sets* $\{B_j^{'}\}$ *at each element* x_i, $i = \overline{1,N}$ *by* (2.4). *Denote these values by* $S(x_1), S(x_2),..., S(x_N)$ *respectively. Compute the value of* $S_{\max}$ *by* (2.5).

Step 2: *Choose from the set* $\{S(x_i)\}$ *such an element* S^* *which is greater than or equal to any other element except* $S_{\max}$.

Step 3: *Do* Step 4 *for* $i = \overline{1,N}$.

Step 4: Compute $\Delta = S^* - S(x_i)$:

If $\Delta < 0$ *then* $\mu(x_i) = \mu_{B_j}(x_i)$;

If $\Delta = 0$ *then compute the value of* $\mu(x_i)$ *by* (2.13);

If $\Delta > 0$ *then compute the value of* k_i *from the equation* $S^* = k_i S(x_i) + (1-k_i) S_{\max}$ *and the value of* $\mu(x_i)$ *by* (2.11) *and* (2.16).

Step 5: *Representation is* $\{\mu(x_1),\ \mu(x_2), ...,\ \mu(x_N)\}$.

IV. ILLUSTRATIONS

Here we give an example of the practical application of the introduced method. Suppose that we have some new project (from an arbitrary area). In the first place the parameterization of the project is carried out, i.e. the components (and their measured values) of the project vector are determined. Some of these parameters can be already determined and their values will be used as they are in the final project decision. To determine the final values of parameters with incomplete initial data it is necessary to invite a group of experts. The task of the project manager (managers) is to obtain the experts' evaluations of each uncertain (fuzzy) parameter and on their basis to adopt the final decision.

The manager offers the group of experts to give a numerical evaluation of each component of the project vector. Suppose that the project suggests the construction of a new hydroelectric power station. Then the following parameters may serve as examples of uncertain data of the project vector:

x_k is the operating costs. They directly depend on the predicted quantity of distributed electric power, on the varying prices of fuel, lubricant, transport, various kinds of maintenance and so on. Hence it can be stated that this parameter possesses a high degree of uncertainty;

x_l is the rate of profit (return) on the equity capital. In all cases this parameter possesses some degree of uncertainty because profit is a predicted value;

x_m is the loan capital share in the total capital amount. In all cases this parameter depends on many variable values such as an interest rate of credit auctions, an interest rate of a bank loan, an interest rate of a bank deposit, a risk factor, specifications of an investment program (e.g. the invested capital cost) and so on. Therefore this parameter contains elements of uncertainty;

x_n is the quantity of distributed electric power. This value is taken from the capital balance, but the balance itself is predictable and besides this value is influenced by the chanding demand as well as by other market factors. Hence this parameter contains elements of uncertainty ; and so on.

After obtaining the subjective evaluations of the experts, i.e. the numerical values of the above-listed parameters of the vector, these values are projected onto the real interval interval [0; b] by means of the normalizing coefficients . Assume that the vector consists of N components and we have m experts. In that case a finite collection of fuzzy sets is formed, which consists of m sets, each of them containing N elements $\{A_j\},\ j=\overline{1,m},\ m=2,3,...$.

Now we must obtain the resulting vector of the components of the considered project. Let us describe how we can do this.

Let the universe X be a finite set $\{x_1, x_2, ..., x_N\}$, $N \ge 1$, the metric lattice $\Psi(X) = \{\mu \mid \mu : X \to [0;b] \subset \mathbb{R}\}$ and the isotone valuation $v(A) = \sum_{i=1}^{N} \mu_A(x_i),\ A \in \Psi(X)$. It is obvious that $v(A \cup B) + v(A \cap B) = v(A) + v(B)$ and $A \subseteq B \Rightarrow v(A) \le v(B)$.

It is easy to see that this isotone valuation v is continuous (see Section 2).

By (2.3) the valuation v determines the following metric:

$$\rho(A,B) = \sum_{i=1}^{N} \left| \mu_A(x_i) - \mu_B(x_i) \right|,\ x_i \in X,\ A,B \in \Psi(X) \quad (4.1)$$

Note that in this case the isotone valuation is the ∑Count, while the metric coincides with the Hamming distance.

Using metric (4.1), the formula for determining the coordination index of a finite collection of fuzzy sets (2.6) takes the form

$$S\{A_j\} = q\left(bN - [(2m+1)/4]^{-1} \sum_{j=1}^{m} \sum_{i=1}^{N} \left| \mu_{A^*}(x_i) - \mu_{A_j}(x_i) \right| \right).\ q > 0 \quad (4.2)$$

Without the loss of generality it can be assumed that the coordination index of a finite collection of fuzzy sets changes in the interval [0;1]. Then by (4.2) the coordination index of the finite collection of fuzzy sets $\{A_j\}$ yields the expression

$$S\{A_j\} = \frac{1}{bN}\left(bN - [(2m+1)/4]^{-1} \sum_{j=1}^{m} \sum_{i=1}^{N} \left| \mu_{A^*}(x_i) - \mu_{A_j}(x_i) \right| \right) \quad (4.3)$$

It is clear that $S_{\max} = 1$.

Now let us evaluate the coordination index of the finite collection of fuzzy sets at each element of the universe. From (4.3) it follows that the coordination index yields the expression

$$S\{A_j\} = \frac{1}{b}\left(b - [(2m+1)/4]^{-1} \sum_{j=1}^{m} \left| \mu_{A^*}(x_i) - \mu_{A_j}(x_i) \right| \right),\ \forall x_i \in X \quad (4.4)$$

By Theorem 2.3 we determine the respective similarity coefficient k_i at each element of the universe:

$$k_i = (1 - S^*) \big/ (1 - S(x_i)). \quad (4.5)$$

It is easy to see that for our case (2.11) takes the form

$$c_i = \frac{1-k_i}{m} \sum_{j=1}^{m} \mu_{B_j^{'}}(x_i)\ . \quad (4.6)$$

and in (2.16)

$$v\left(B_j^{'}\right)=\mu_{B_j^{'}}\left(x_i\right),\quad \forall x_i\in X\ . \qquad (4.7)$$

We give the following concrete numerical example. Let the universe X be a finite set $\left\{x_1,x_2,...,x_7\right\}$ and $\Psi\left(X\right)=\left\{\mu\mid\mu:X\to\left[0;10\right]\right\}$ be a metric lattice with continuous isotone valuation $v\left(B\right)=\sum_{i=1}^{7}\mu_B\left(x_i\right)$. Assume that we have a group of five experts, each of whom evaluated numerically seven uncertain parameters (components) of the project vector *B. Then we project these values onto the real interval* [0; 10]. As a result we obtain the finite collection of fuzzy sets $\left\{B_j\right\},\ j=\overline{1,5}$:

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
B_1	1.0	8.0	6.4	4.8	2.0	7.2	3.0
B_2	1.5	7.4	6.5	3.2	4.0	8.1	3.4
B_3	2.1	8.2	7.0	5.0	3.0	6.8	3.0
B_4	2.5	9.0	6.7	5.5	2.6	8.2	5.0
B_5	1.9	9.5	6.8	3.5	2.5	8.5	5.5

(4.8)

and by Definition 2.2 determine its increasing shuffling $\left\{B_j^{'}\right\}$:

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$B_1^{'}$	1.0	7.4	6.4	3.2	2.0	6.8	3.0
$B_2^{'}$	1.5	8.0	6.5	3.5	2.5	7.2	3.0
$B_3^{'}$	1.9	8.2	6.7	4.8	2.6	8.1	3.4
$B_4^{'}$	2.1	9.0	6.8	5.0	3.0	8.2	5.0
$B_5^{'}$	2.5	9.5	7.0	5.5	4.0	8.5	5.5

(4.9)

Simplify (4.4) for our case:

$$S\{B_j^{'}\}=0.1\times(10-0.5\times\sum_{j=1}^{5}|\,\mu_{B^*}(x_i)-\mu_{B_j^{'}}(x_i)\,|),\quad \forall x_i\in X\ .$$

Note that here the representative B^* is easily determined by Theorem 2.1. Using the last expression we calculate the values of the coordination index of the increasing shuffling at each element of the universe:

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$S\{x_i\}$	0.895	0.845	0.955	0.810	0.875	0.865	0.775

(4.10)

The coordination index does not reach its maximal value at every element of the universe. Denote the greatest value by S^* = 0.955.

By (4.5) and (4.10) we obtain $k_1\ \approx 0.4286$, and. by (4.6) and (4.9) $c_1\approx 1.0286$. Now by (2.16) and (4.9) with (4.7) taken into account we compute $\mu(x_1)\ \approx\ 1.8257$. The fuzzy aggregation at all other elements of the universe is obtained analogously. As a result we have

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$k_i\approx$	0.4286	0.2903	1.0	0.2368	0.36	0.(3)	0.2
$c_i\approx$	1.0286	5.9755	–	3.3579	1.8048	5.1733	3.184
$\mu(x_i)\approx$	1.8257	8.4087	6.65	4.4340	2.7588	7.6518	3.848

(4.11)

We know that all experts have the same qualification. We may find out which of them had the highest concentration of evaluation effort in the course of evaluations. For this we have to calculate a standard deviation of the evaluations from the result of the group decision-making for each expert:

$$\delta_j=(\sum_{i=1}^{7}|\,\mu(x_i)-\mu_{B_j}(x_i)\,|)/7.$$

Using (4.8), (4.9) and (4.11), from the last expression we obtain

No.	1	2	3	4	5
$\delta_j\approx$	0.5584	0.6937	0.4771	0.6058	0.7155

,

whence we conclude that expert no. 3 had the highest concentration of evaluation effort.

V. Conclusion

We consider the approaches to making decisions for control problems in nonstandard situations with a lack of the previous experience and incomplete knowledge of the considered problem. In such cases we usually cannot do without expert evaluations which lead to the process of group decision-making, and it becomes necessary to solve a problem of alternatives aggregation. It has been proposed to solve such problems by means of fuzzy sets. The material of the paper is divided into two parts. In Part 1, an approach is proposed for the processing of quantitative expert evaluations which are used in the group decision-making. The approach is based on the coordination index and the similarity of finite collections of fuzzy sets and takes into account the specific character of the fuzzy aggregation operator. The approach is discussed in full detail and its algorithm is presented. An example of the application of the proposed method is given.

The approach to making decisions for control problems is considered in nonstandard situations, in which there is a lack of the previous experience and the available information is uncertain and/or incomplete and/or weakly structured. In the framework of the proposed approach, quantitative expert evaluations are processed by using a special mathematical apparatus that employs the coordination index, the similarity of finite collections of fuzzy sets and takes into account the specific character of a fuzzy aggregation operator. The coordination index of a finite collection of fuzzy sets introduced in a new way forms the basis of the proposed approach. Setting aside the question of independence and completeness of the introduced postulates (Definition 2.3), by constructing the corresponding example [7] we prove that this axiomatics is consistent. The proposed approach is discussed in full detail and its algorithm is presented. An example of the application of the proposed method is given.

ISBN: 978-1-61804-227-9

The prospects of the obtained results are briefly as follows. By virtue of (2.7) the proposed fuzzy aggregation method makes it possible to increase the coordination degree of experts' evaluations by $nS^* \Big/ \sum_{i=1}^{n} S_i$ times (here $X = \{x_1, x_2, \ldots, x_n\}$, S_i are the coordination indices at each element of the universe, S^* is a maximal coordination index among all one-element fuzzy sets).

The method may also be helpful for a group of experts in the implementation of a project. Given a project with its vector of goals, the problem is to determine the coordination degree of evaluations made by contesters claiming to join the project group. The proposed approach enables one to calculate this degree and collect the most efficient project group.

REFERENCES

[1] A. Averkin, I. Batirshin, A. Blishun. V. Silov, V. Tarasov, Fuzzy sets in the Models of Control and Artificial Intelligence (Nauka Press, Moscow, 1986) (in Russian).

[2] D. Dubois, H. Prade, Possibility theory, Plenum, New York, 1988

[3] M.Meskon , M.Albert, F.Kheduri - Management – 3rd Edition: translation from English- M.: LLC "I.D. Williams", 2007.-p.602.

[4] P.Schoderbek – Management Systems – 2nd ed – New York: Wiley, 1971, p.124.

[5] H.Simon - The Sciences of the Artificial. The MIT Press. ISBN:0-262-69191-4; 978-0-262-6 9191-8; 1996.

[6] H.Simon, D.Smithburg,V.A. Thompson- Public Administration. Alfred A. Knopf:New York, 1961, p.493.

[7] T. Tsabadze, The coordination index of a finite collection of fuzzy sets, Fuzzy Sets and Systems 107 (1999) 177-185.

[8] T. Tsabadze, A method for fuzzy aggregation based on grouped expert evaluations, Fuzzy Sets and Systems 157 (2006) 1346-1361.

[9] J. Vaníček, I. Vrana and S. Aly, Fuzzy aggregation and averaging for group decision making: A generalization and survey. Knowledge-Based Systems 22 (2009) 79-84.

[10] L. A. Zadeh, Fuzzy sets, *Inform. Control* 8, No. 3 (1965), 338–353.

Multi-level Image Classification Using Fuzzy Petri Net

Marina Ivasic-Kos, Slobodan Ribaric, Ivo Ipsic

Abstract**—** For a multi-level image classification, a knowledge representation scheme based on Fuzzy Petri Net with fuzzy inference algorithms is used. A simple graphical Petri net notation and a well-defined semantics displaying the process of reasoning through inference trees are used for visualization of the knowledge base and explanations of derived conclusion. Used knowledge representation formalism has the ability to show a probability of concepts and relations. The procedures of image multi-level classification using fuzzy recognition and inheritance algorithms on a knowledge representation scheme, as well as experimental results of image semantic interpretation, are presented.

Keywords**—** Fuzzy Petri Net, Image Classification, Knowledge Representation

I. Introduction

Digital images have become unavoidable in the life of modern people. The number of digital images on a specialized web site counts in millions, and private and business collections are rapidly increasing. A large number of images is becoming a current problem for search and retrieval, as well as for organizing and storing.

It is believed that we could retrieve and arrange images simply if they can be automatically annotate and describe with words that are used in an intuitive image search. People for the interpretation of the images usually do not use objects that appear in the image, but the broader context that arises from the relation between these objects and scenes. However, the content of images is generally difficult to typify, and sometimes is not even easy to describe the images in words that will meet different requirements and needs.

To interpret image as people do, feature extraction and recognition of objects are not sufficient. A multi-level classification, semantic modeling and representation of knowledge that is specific to the application domain are necessary. Thus, the process of multi-level image classification should include low-level image features extraction, then learning and implementation of a model that maps image features to classes that can be recognized in the image and a knowledge acquisition to determine the parent classes. The amount of knowledge that is needed for the classification of images increases with the semantic level of concepts used to interpret the images.

For solving the problem of object recognition many different approaches has been used. A recent survey and research made in the field can be found in [1]. For multi-level image classification and interpretation several different approaches that use machine learning techniques or models for knowledge representation and reasoning were proposed in recent years and hereafter we will mention some of them.

In [2] a Multi-Level Image Sampling and Transformation methodology (MIST) is described that uses a neural network as a classifier and symbolic rules from the knowledge base in order to semantically interpret a new image.

A hierarchical model for generating words that correspond to class labels is proposed in [3]. The model is inspired by the Hofmann's hierarchical clustering model and a model of soft clustering.

In [4] a SVM classifier is used for learning the elementary classes of natural scenes, which are then using probabilistic model linked in concepts of a higher semantic level, such as "beach".

To view the perceptual and semantic information of multimedia content a semantic network is used in [5].

One of the early works that uses the ontology for the semantic description of the image content and descriptive logic for verification of the classification results is [6].

To explore the ontology of words that is used for image interpretation and annotation [7] have used WordNet. This idea is further extended in [8]. The authors intend to create public image ontology, the ImageNet, with aim to illustrate each of the concepts from the WordNet ontology with 500-1000 images.

Within the project aceMedia, [9] combine ontology with fuzzy logic to generate concepts from beach domain with appropriate reliability. In [10], the same group of authors have used the SVM classifier and inference engine that supports fuzzy descriptive logic and in [11] a combination of different classifiers for learning concepts and fuzzy spatial relationships. Authors have reported that environment used by the ontology is shown to be incompatible with that of fuzzy reasoning engines.

In this paper, for multi-level image classification, a knowledge representation scheme based on Fuzzy Petri Net is used. A knowledge formalism and inference engine is briefly presented in chapter two. A proposal of multi-level image classification is given in the third chapter. An example of

M. Ivasic-Kos is with the Department of Informatics, University of Rijeka, Rijeka, Croatia (corresponding author e-mail: marinai@uniri.hr).

S. Ribaric is with the Department of Electronics, Microelectronics, Computer and Intelligent Systems, Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia (e-mail: slobodan.ribaric@zemris.fer.hr).

I. Ipsic is with the Department of Informatics, University of Rijeka, Rijeka, Croatia (e-mail: ivoi@uniri.hr).

ISBN: 978-1-61804-227-9

knowledge base, which relates to outdoor domain, is shown in chapter four. Furthermore, a description of derived conclusion using inheritance and recognition trees, as well as experimental results is given in chapter five, six and seven, respectively.

II. Knowledge Formalization

For multi-level image classification a knowledge representation scheme based on Fuzzy Petri Net, named KRFPN, [12] is used.

A. Definition of Knowledge Representation Scheme

The KRFPN scheme is defined as 13-tuple:

KRFPN = (P, T, I, O, M, Ω, μ, f, c, α, β, λ, Con), (1)

where:

$P = \{p_1, p_2 ... p_n\}$, $n \in \mathbb{N}$ is a set of places,
$T = \{t_1, t_2 ... t_m\}$, $m \in \mathbb{N}$ is a set of transitions,
$I: T \to P^{\infty}$, is an input function,
$O: T \to P^{\infty}$, is an output function,
$M = \{m_1, m_2 \ldots m_r\}$, $1 \le r < \infty$, is a set of tokens,
$\Omega: P \to \wp(M)$, is a tokens' distribution within places,
$\mu: P \to N$, marking of places,
$f: T \to [0, 1]$, the degree of truth of the transitions,
$c: M \to [0, 1]$, the degree of truth of the token,
$\alpha: P \to D$, maps place from set P to concept from set D,
$\beta: T \to \Sigma$, maps transition from set T to relation in set Σ,
$\lambda \in [0, 1]$, threshold value related to transitions firing,
$Con \subseteq (D \times D) \cup (\Sigma \times \Sigma)$, is a set of pairs of mutually contradictory relations or concepts.

The KRFPN can be represented by a direct graph containing two types of nodes: places and transitions. Graphically, places $p_i \in P$ are represented by circles and transitions $t_j \in T$ by bars. The relationships, based on input and output functions are represented by directed arcs. In a semantic sense, each place from set P corresponds to a concept from set D and any transitions from set T to relation from set Σ (Fig. 1).

A dot in a place represents token $m_1 \in M$, and the place that contains one or more tokens is called a marked place. Tokens give dynamic features to the net and define its execution by firing an enabled transition. The transition is enabled when every input place of transition is marked, i.e. if each of the input places of the transition has at least one token. Moreover, if threshold value λ that defines the sensitivity of the knowledge base is set, truth value $c(m_1)$ of each token must exceed the value of λ if the transition would be enabled.

An enabled transition t_j can be fired. By firing, a token moves from all its input places $I(t_j)$ to the corresponding output places $O(t_j)$. In Fig. 1 there is only one input place for transition t_j, $I(t_j) = p_i$ and only one output place $O(t_j) = p_k$. After transition firing, a new token value $c(m_2)$ is obtained as $c(m_1) * f(t_j)$ in the output place (Fig. 2). Values $c(m_1)$ and $f(t_j)$ are degrees of truth assigned to token at the input place $p_i \in I(t_j)$ and transition $t_j \in T$, respectively.

Value of $c(m_1), f(t_j) \in [0,1]$, can be expressed by truth scales where 0 means «no true» and 1 «always true» [13]. Semantically, value $c(m_1)$ express the degree of uncertainty of joining a particular concept form set D to place p_i, and value $f(t_j)$ the degree of uncertainty of links between relationship

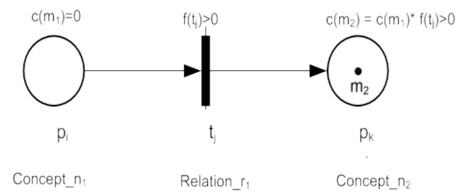


Fig. 2 Fuzzy Petri net formalism with associated semantic meaning

from a set Σ and a transition t_j.

B. Inference Engine

Inference engine of KRFPN scheme consists of three automated reasoning processes: fuzzy inheritance, fuzzy recognition and fuzzy intersection.

All inference processes are based on dynamic properties of the network, and are graphically shown by the inheritance or the recognition tree. The steps of all algorithms are given in [12].

This paper describes the use of fuzzy inheritance and fuzzy recognition algorithms for multi level image classification.

III. Image Semantic Categories

The goal is to classify images as much as possible closer to the semantic concepts that people use when interpreting these images. Therefore, proposed multi-level image classification includes classes from four semantic levels – an elementary class, a generalization class, a derived class and a scene class. Elementary classes correspond to object which were directly identified in the image like "train", "airplane" or "sky".

Other semantic class categories are used for the interpretation of images on higher-level and are defined according to expert knowledge. Generalization classes include classes which were created by generalizing objects recognized in the image or in the case of high-level generalization by generalizing already generalized classes like: "airplane" (elementary class) - "vehicle" (generalization of elementary

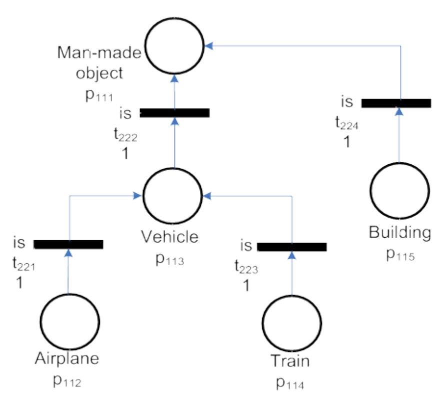


Fig. 3 A part of knowledge base displaying class generalization

class) – “man-made object”(high-level generalization), Fig 3.

Same abstract classes that are “common” to human interpretation of some objects like “winter” for “snow” can be described by derived classes.

Scene classes are used to represent the semantics of the whole image like “mountain view”, “natural scene”, “outdoor”.

IV. Knowledge Base Definition

A. *Definition of Concepts*

According to the KRFPN scheme, classes from all semantic levels are elements of a set D, where D = D1 ∪ D2 ∪ D3. A subset D1 includes generalized classes (e.g. set GC ={Outdoor Scenes, Natural Scenes, Man-made Objects, Landscape, Vehicles, Wildlife, …}), scene classes (e.g. set SC ={Seaside, Inland, Sea, Underwater, Space, Airplane Scene, Train Scene, …}) and related derived or abstract classes (e.g. set AC ={Summer, …}) defined according the expert knowledge, thus $D_1 = SC \cup GC \cup AC$.

A subset D2 is used in case of some special instance of classes of interest or for instance of unknown class X that should be determinate.

A subset D3 represents class attributes and in this experiment consists of elementary classes like C = {Airplane, Train, Shuttle, Building, Road, Grass, Ground, Cloud, Sky, Coral, Dolphin, Bird, Lion, Mountain, …} that are according to modified Bayesian rule selected as attributes of scenes. It is assumed that a scene may contain several characteristic elementary classes, so instead of choosing an attribute with a maximum posterior probability, all those elementary classes with a posterior probability $P(SC_i|C_k), \forall_k C_k$ exceeding the marginal value ε for a given scene SC_i $\forall_i$ are selected:

$$M_k = \{m_{ki} : \arg_i P(C_k|SC_i)P(SC_i) \geq \varepsilon\}. \qquad (2)$$

A. *Definition of Relations*

Relations from a set Σ are defined according to expert knowledge.

The set Σ is a union of sets $\Sigma_1 \cup \Sigma_2 \cup \Sigma_3$, where subset Σ_1 is set of hierarchical relations (e.g. Σ_1= {is, is part of}, Σ_2 is a set of relations between class and values of its attributes (in these case elementary classes) from set D_3 (e.g. Σ_2 = {consist}) and subset Σ_3 is a set of spatial and pseudo-spatial relations defined by a spatial location of objects and by co-occurrence of objects in real scenes, respectively, like Σ_3 = {is below, is above, occurs with, occurs not with, , …}.

A. *Definition of Relations Truth Value*

For the relations from the set Σ_1 that model the class inheritance, degree of truth is set to 1 because any exceptions, if exist, can be modeled using a set of contradictions. Truth value of the relations, linking the elementary class and derived class, is defined according to the experts’ knowledge. Truth value of relation whether from the sets Σ_2 and Σ_3 is computed using data in the training set.

Truth value attached to the relation between attributes and classes was determined using discriminate function $d_i(C_k)$ defined separately for each attribute C_k. It is assumed that attributes are independent. The model is adjusted on a learning set and effectiveness of selected attributes is evaluative on the test set. Decision rule is:

$$C_k \,\epsilon\, SC_i \overset{iff}{\Leftrightarrow} j = \underset{i}{\arg\max}\, d_i(C_k), \;\; \forall_i SC_i, \forall_k C_k$$
$$d_i(C_k) = \frac{P(C_k|SC_i)P(SC_i)}{\sum_{j=1}^{n} P(C_k|SC_j)P(SC_j)} \qquad (3)$$

Moreover, to give greater importance to attributes with more contribution on the classification results, attribute weights are estimated by misclassification error (MCE criterion):

$$w(C_i) = 1 + \frac{1 - MCE(C_i)}{\sum_k MCE(C_k)}, \;\; \forall_k C_k, \quad 1 \leq k \leq |C|. \qquad (4)$$

In Fig. 4 a part of knowledge base is presented, showing relations among particular scene class and appropriate elementary classes defined by the former procedure. For example, the degree of truth of relation between a particular

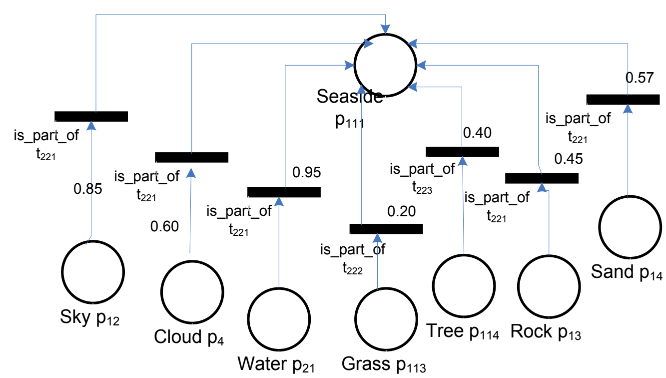


Fig. 4 Relations among scene ‘Seaside’ and appropriate elementary classes

class “Seaside” and its attribute, an elementary class “Water” is set to 0.95.

Analyzing a mutual occurrence of classes C_j, C_i a truth value of pseudo-spatial relations can be formally defined as:

$$P(C_j|C_i) = \frac{P(C_j \cap C_i)}{P(C_i)}, \qquad i \neq j \qquad (5)$$

The spatial and pseudo-spatial relationships can be used to validate and adjust the results of classification. For instance, having detected the “airplane” in the image with high probability, detection of “lion” on the same image is unlikely. Also, grass often appears under the sky, the sky above the mountains, etc.

ISBN: 978-1-61804-227-9

V. Scene Classification Using a Fuzzy Recognition Algorithm

For a task of scene classification of a new, unknown image, fuzzy recognition algorithm on inverse KRFPN scheme is used. The inverse KRFPN scheme is obtained by replacing the input and output functions of KRFPN scheme and is denoted as -KRFPN. The procedure of fuzzy recognition finds the class whose properties best match given set of attributes and relations.

Assumption is that unknown image is segmented and that low-level image features are obtained. Using some classification method as Naive Bayes, each image segment is classified in one of elementary classes according a maximum posterior probability (c_{MAP}).

The basic assumption is that attributes (feature vector components) within the class are mutually independent, so that applies:

$$P(x_1, x_2, \dots, x_n \mid C_j) \approx \prod_{j=1}^{n} P(x_i \mid C_j). \tag{6}$$

Based on the Bayes' theorem and taking into account that $P(x) = const.$ and judging $P(C_i)$, $\forall C_i \in C$ on the basis of data in a learning set for each attribute value x_j of new data occurrence $\mathbf{x}^{new}$, a classification results is determined by:

$$c_{MAP} = \underset{C_i \in C}{\operatorname{argmax}} P(x^{new} \mid C_i)\, P(C_i) \tag{7}$$

The results of the classification of each segment and assessment of a posterior probability $P(C_j \mid x_i)$ are entry of the Petri net used for further classification on higher semantic level.

Thus, a set of obtained elementary classes C_i are treated as attributes of an unknown scene class X that are mapped to places $\{p_1, p_2, \dots, p_k\}$ if a function $\alpha^{-1}: C_i \to p_k$ is defined. A token value $c(m_k)$ of each place corresponds to obtain posterior probability of the appropriate elementary class C_i mapped to that place.

The initial token distribution will be a root node π_0 of the recognition tree (Fig. 5).

Fig. 5 shows corresponding recognition trees in -KRFPN scheme with enabled transition starting from the root node. Nodes of the recognition tree have a form $(p_j, c(m_l))$ $j = 1, 2, \dots, n$, $l = 1, 2, \dots, r$, $0 \le r \le |M|$, where $c(m_l)$ is a value of token m_l in place p_j. Arcs of recognition tree are marked with value $f(t_j)$ and label of a transition $t_j \in T$ whose firing creates new nodes linked to scene classes.

The following describes the procedure for the classification of elementary classes in the scene classes using fuzzy recognition algorithm that matches the recognition trees shown in Fig. 5.

For instance, if obtained classification results are elementary classes that exist in knowledge base with corresponding degree of truth: (cloud {0.5}, rock ({0.4}), sand ({0.8}), water ({0.8}) than using function α^{-1} initially marked places are determinate (α^{-1}(cloud)=p4, α^{-1}(rock)=p13, α^{-1}(sand)=p14, α^{-1} (water)=p21). According to initially marked places and corresponding degree of truth, four recognition trees $\pi^i, i = 1,2,..4$ with root node $\pi_0{}^i, i = 1,2,..4$ will be formed (Fig. 5):

$$\pi_0^1(\emptyset, .. \emptyset, \{p4, 0.5\}, \emptyset, \dots, \emptyset), \pi_0^2(\emptyset, .. \emptyset, \{p13, 0.4\}, \emptyset, .. \emptyset,),$$
$$\pi_0^3(\emptyset, .. \emptyset, \{p14, 0.8\}, \emptyset, .. \emptyset,), \pi_0^4(\emptyset, .. \emptyset, \{p21, 0.8\}, \emptyset, .. \emptyset,)$$

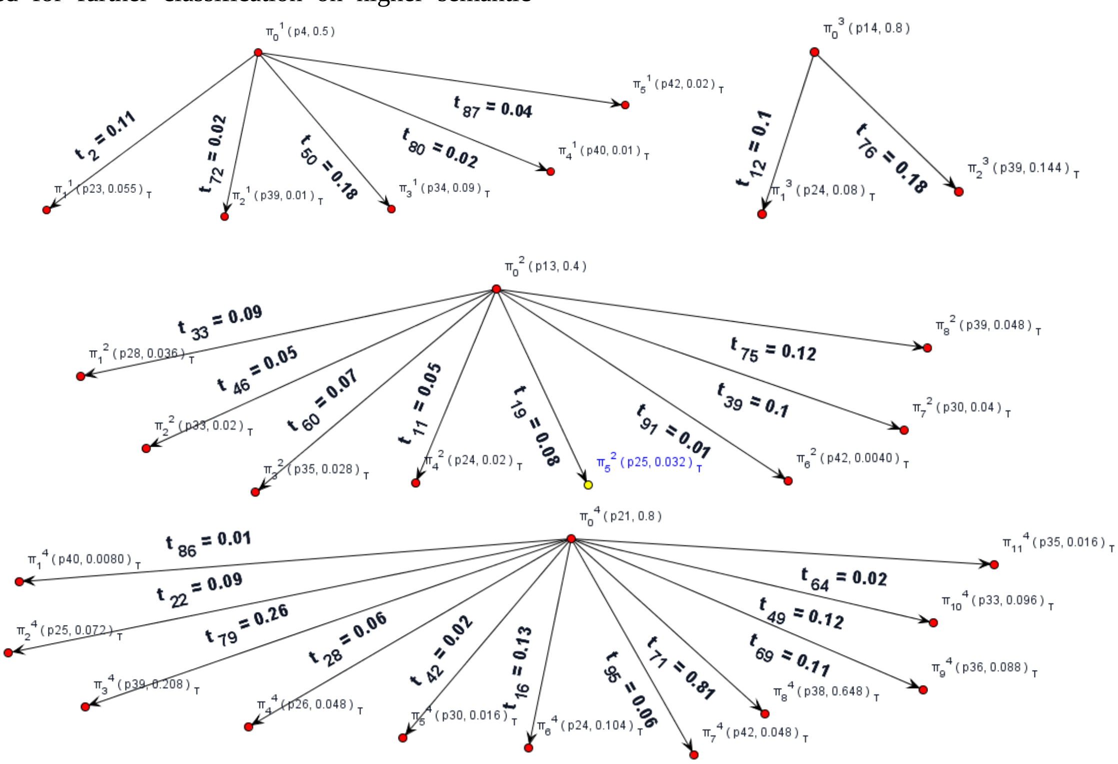


Fig 1 Recognition trees with enabled transitions for root nodes that match initially obtained elementary classes

ISBN: 978-1-61804-227-9

By firing of enabled transitions on - KRFPN scheme, new nodes on the following higher level of recognition tree are created and appropriate values of tokens are obtained:

$$c(m_{k+1}) = c(m_k) * f(t_k) = P(C_i|\mathbf{x}) * w_i(C_i)P(C_i|SC_l)P(SC_l) \quad (8)$$

where t_k is arc between concepts C_i and SC_l, $P(C_i|\mathbf{x})$ is a posterior probability of segment classification to elementary class C_i.

Accordingly, total sum of all p nodes $\pi_i{}^k, i = 1,2,..p \leq n$ $k = 1,2,..b \leq |M|$ in all b recognition trees (for this example b=4 and p=$\sum_{k=1}^{4} p^k = 5 + 8 + 2 + 11 = 28$), excluding the root node $\pi_0{}^k$, is computed as:

$$Z = \sum_{k=1}^{b} \sum_{i=1}^{p^k} \pi_i^k. \quad (9)$$

In this example a total sum is:

$Z = \sum_{k=1}^{4} \sum_{i=1}^{p^k} \pi_i^k = \sum_{i=1}^{5} \pi_i^1 + \sum_{i=1}^{8} \pi_i^2 + \sum_{i=1}^{2} \pi_i^3 + \sum_{i=1}^{11} \pi_i^4$ = (∅…∅, {p23, 0.055}, {p24, 0.204}, {p25, 0.104}, {p26, 0.048}, ∅, {p28, 0.036}, ∅, {p30, 0.056}, ∅, ∅, {p33, 0.116}, {p34, 0.09}, {p35, 0.044}, {p36, 0.088}, ∅,{p38, 0.648}, {p39, 0.41}, {p40, 0.018}, ∅, {p42, 0.072}, ∅, ... ∅).

Then, a set of indices of elements with a highest sum $Z = (Z_1, Z_2, \dots, Z_n)$ among all of the nodes in all recognition trees is selected:

$$I^* = \{i^* : \arg \max_{i=1,\dots,n} \{Z_i\}\} \quad (10)$$

A scene class assigned to a place with max argument $p_i : i \in I^*$ is chosen as the best match for a given set of elementary classes.

In this example, a set of max argument is $I^* = \{38\}$ and a scene class chosen as the best match for a give set of attributes is one that is assigned to place with max argument, α (p38) =‘Seaside’.

Obtained classes can be used as root nodes for next recognition process that will infer classes from higher semantic levels either because they are directly linked with the classes or may be inferred by means of classes (parents) at a higher level of hierarchy.

VI. Class Generalization Using Fuzzy Inheritance

To display the properties of the concept and its relations, fuzzy inheritance algorithm on the KRFPN scheme can be used. The fuzzy inheritance algorithm determines attributes of a classes $d_j \in D$, first locally and then at higher hierarchical levels than the classes d_j it selves. During the process of inheritance for a given $k \in \mathbb{N}$, a final tree of inheritance at the most k+1 level is constructed. As the class of interest can be at different levels of abstraction, whether at the level of the elementary class or the scene class, a key feature of inheritance algorithm is that allows the representation of knowledge at different levels of abstraction.

For a given class that exists in the database, the appropriate place is determined by the function $\alpha^{-1}(d_j) = p_k,\ d_j \in D$.

According to the initially marked place and appropriate token value, the initial token distribution is created $\Omega_0 = \pi_0 = (\emptyset, \emptyset, \dots, \{p_k, c(m_k)\}, \dots, \emptyset)$, which represents the root node of inheritance tree. Token value $c(m_k)$ can be set to 1 or to the value obtained by recognition algorithm.

The inheritance tree is formed by firing the enabled transitions until the condition for stopping the algorithm is satisfied or the desired depth of inheritance tree reached.

Below, a Fig. 6 shows a 1-level inheritance tree of the KRFPN scheme for one of scene classes, a "Seaside", where $\alpha^{-1}('Seaside') = p_{39}$, $c(m_1) = 1$ and the corresponding root node is $\pi_0(\emptyset, \dots \emptyset, \{p39, 1.0\}, \emptyset, \dots, \emptyset)$.

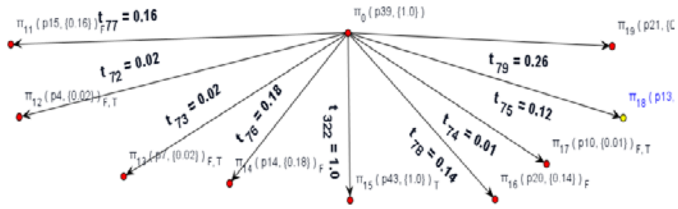


Fig. 6 The inheritance tree for a concept Seaside

Fig. 7 shows inheritance paths formed after semantic interpretation of nodes and arch displayed in Fig. 6 describing, semantically, attributes and parent classes of class “Seaside”.

Inheritance inference procedure for concept Seaside gives following conclusior

q.	Inheritance path interpretation	Associated trut
	Seaside contains cloud.	Minimally true
	Seaside contains grass.	Minimally true
	Seaside contains mountain.	Minimally true
	Seaside contains rock.	Slightly true
	Seaside contains sand.	Slightly true
	Seaside contains sky.	Slightly true
	Seaside contains trees.	Slightly true
	Seaside contains water.	Slightly true
	Seaside is Obala.	Always true

Fig. 7 Statements of inheritance for concept 'Seaside'

VII. Experimental Results

To demonstrate a model of hierarchical image classification using the KRFPN scheme, we have used a part of Corel image dataset [14].

Images were automatically segmented based on visual similarity of pixels using the Normalized Cut algorithm [16], so segments do not fully correspond to objects. Every segmented region of each image is more precisely

ISBN: 978-1-61804-227-9

characterized by a set of 16 features based on color, position, size and shape of the region [14].

Also, each image segment of interest was manually annotated with first keyword from a set of corresponding keywords provided by [16] and used as ground truth for the training model. Vocabulary used to denote the segments have 28 words related to natural and artificial objects such as 'airplane', 'bird', etc. and landscape like 'ground', 'sky', etc.

The data set used for an experiment consists of 3960 segments divided into training and testing subsets by 10-fold cross validation with 20% of observations for holdout cross-validation.

We have used the Naïve Bayes classification algorithm to classify image segment into elementary classes. The results of automatic classification of image segments are compared with ground truth, so the precision and recall measures are calculated, Fig. 8.

A recall is the ratio of correctly predicted classes and all classes for the image (ground-truth), while a precision is the ratio of correctly predicted classes, and total number of suggested classes.

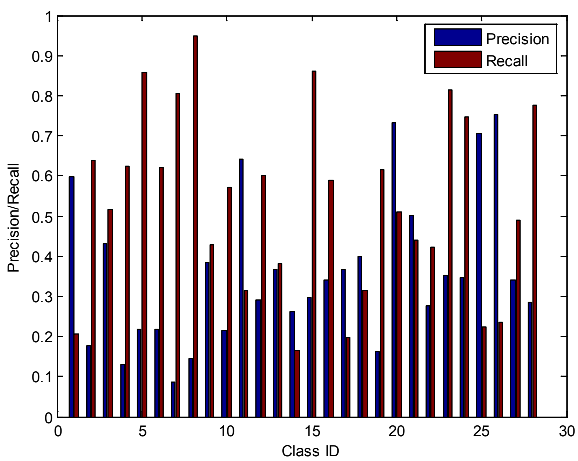


Fig. 8 A precision/recall graph for classification of image segments into elementary classes (displayed by ID)

The results of low-level image feature classification depend on the quality of segmentation, so when image has a lot of segments and when object is over segmented, the results can include labels that do not correspond to the context of image. Then, using the facts from the knowledge base, the obtained results are analyzed with the fuzzy inheritance algorithms in order to purify the classification results from class labels that do not match the contents of the image.

Afterwards, based on elementary classes obtained as classification results and knowledge developed for particular domain, an automatic image classification on higher semantic level can be performed following the fuzzy recognition algorithm. Also, the inheritance algorithm can be used to display the properties of the class and its relations with the parent classes. In Table 1 some examples of results of a multi-level image classification are indicated including results of low level image classification (the 1. row below each image) and image classification using proposed knowledge scheme (the 2. row below each image).

Table 1: Examples of scene classification

		
'train', 'tracks', 'sky'	'grass', 'tiger'	'water', 'sand', 'sky', 'road'
'Vehicle', 'Man-Made Object', 'Outdoor'	'Wildcat', 'Wildlife', 'Natural Scenes', 'Outdoor Scene'	'Coast', 'Landscape', 'Natural Scenes', 'Outdoor Scene'

VIII. CONCLUSION

The aim of this paper is to present a model for multi-level image classification using KRFPN formalism that uses fuzzy Petri Nets graphical notation. A precise mathematical model of Petri nets and inference algorithms, with finite recognition and inheritance inference trees, can be used to present and analyze, whether the relationships between attributes and class or between classes at a higher level of abstraction.

The complexity of the algorithm is O (nm) where n is the number of places and m number of transitions in KRFPN scheme.

A hierarchical organization of KRFPN scheme enables explanation of image classification results obtained by some other classification method, such as Naive Bayes classifier.

Furthermore, an important property of the KRFPN formalism is the ability to show the uncertain knowledge using probability or reliability of the concept and relation.

This research is limited to a domain of outdoor scenes and the knowledge base includes knowledge that is relevant to the domain. But, the methodology of acquiring knowledge and reasoning in KRFPN scheme is expandable and adaptable to the acquisition of new knowledge of a particular domain.

REFERENCES

[1] Datta, R., Joshi, D., Li, J. 2008. “Image Retrieval: Ideas, Influences, and Trends of the New Age”, ACM Transactions on Computing Surveys, vol. 20, pp. 1-60, April 2008.

[2] Michalski, R.S., Zhang, Q., Maloof, M.A., Bloedorn, E., 1996. “The MIST Methodology and its Application to Natural Scene Interpretation”, Proc. of the Image Understanding Workshop, Palm Springs, CA, Feb. 1996.

[3] Barnard, K., Duygulu, P., Forsyth, D., Freitas, N. de, Blei, D. M., Jordan, M. I., 2003. “Matching words and pictures,” Journal of Machine Learning Research vol. 3, pp. 1107–1135, 2003.

[4] Fan, J., Gao, Y., Luo, H., Jain, R., 2008. “Mining Multilevel Image Semantics via Hierarchical Classification”, IEEE Transactions on Multimedia, vol. 10, 2008, pp. 167-187.

[5] Benitez, A.B., Smith, J. R., Chang, S.F., 2000. “MediaNet: A Multimedia Information Network for Knowledge Representation”, Proc. IS&T/SPIE, v. 4210, MA, November 2000.

[6] Schreiber, A. T. G., Dubbeldam, B., Wielemaker, J., Wielinga, B., 2001. “Ontology-based photo annotation,” IEEE Intelligent Systems, vol. 16, no. 3, pp. 66–74, 2001.

[7] Srikanth, M., Varner, J., Bowden, M., Moldovan, D., 2005. “Exploiting ontologies for automatic image annotation”, Proc. SIGIR’05, 2005, p. 552–558.

ISBN: 978-1-61804-227-9

[8] Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L., 2009. "ImageNet: A Large-Scale Hierarchical Image Database".

[9] Papadopoulos, G.T., Mylonas, P., Mezaris, V., Avrithis, Y., Kompatsiaris, I., 2006. "Knowledge-Assisted Image Analysis Based on Context and Spatial Optimization", Int. Journal on Semantic Web and Information Systems, vol. 2, no. 3, pp. 17-36. July-Sept. 2006.

[10] Mezaris, V., Papadopoulos, G.T., Briassouli, A., Kompatsiaris, I., Strintzis, M.G., 2009. "Semantic Video Analysis and Understanding", chapter in "Encyclopedia of Information Science and Technology", Second Edition, Mehdi Khosrow-Pour, 2009, ebook.

[11] Athanasiadis, T. et al. 2009. "Integrating Image Segmentation and Classification for Fuzzy Knowledge-based Multimedia", Proc. MMM2009, France, 2009.

[12] Ribarić, S., Pavešić, N., 2009. "Inference Procedures for Fuzzy Knowledge Representation Scheme", Applied Artificial Intelligence, vol. 23, January 2009, pp. 16-43.

[13] Chen, S.M., Ke, J.S., Chang, J.F., 1990. "Knowledge Representation Using Fuzzy Petri Nets", IEEE Transactions on Knowledge and Data Engineering, vol. 2, 1990, pp. 311-319.

[14] Duygulu, P., Barnard, K., Freitas, J.F.G. de, Forsyth, D. A., 2002. "Object recognition as machine translation: Learning a lexicon for a fixed image vocabulary", Proc. ECCV 2002, UK, May 2002, pp. 97–112.

[15] Carbonetto, P., Freitas, N. de, Barnard, K., 2004. "A Statistical Model for General Contextual Object Recognition", Proc. ECCV 2004, Czech Republic, May 2004, pp. 350-362.

[16] Shi, J., Malik, J, 2000. "Normalized cuts and image segmentation", IEEE Trans. PAMI, vol. 22, no. 8, pp. 888–905, 2000.

Sensorless sliding mode control of induction motor using fuzzy logic Luenberger observer

A. Bennassar, A. Abbou, M. Akherraz, and M. Barara

Abstract**—**Many industrial applications require high performance speed sensorless operation and demand new control schemes in order to obtain fast dynamic response. In this paper, we present a speed sensorless sliding mode control (SMC) of induction motor (IM). The sliding mode control is a powerful tool to reject disturbances. However, the chattering phenomenon presents a major drawback for variable structure systems. To decrease this problem, a saturation function is used to limit chattering effects. A Luenberger observer based on fuzzy logic adaptation mechanism is designed for speed estimation. Numerical simulation results of the proposed scheme illustrate the good performance of sensorless induction motor and the robustness against load torque disturbances.

Keywords**—**Fuzzy logic control, induction motor, Luenberger observer, sliding mode control.

I. Introduction

THE induction motor is one of the most widely used machines in various industrial applications due to its high reliability, relatively low cost, and modest maintenance requirements [1]. Many industrial applications require high dynamic performances and robustness to different perturbations. Thus, the robust control algorithm is desirable in stabilization and tracking trajectories. The variable structure control can offer a good insensitivity to parameter variation, external disturbances rejection and fast dynamics [2]-[3].

The sliding mode control is a type of variable structure system characterized by high simplicity and robustness against insensitivity to parameters variation and disturbances. This approach utilizes discontinuous control laws to drive the system state trajectory onto a sliding or switching surface in the state space. The dynamic of the system while in sliding mode is insensitive to model uncertainties and disturbances [4]. However, the discontinuous control presents a major drawback presented in chattering phenomenon. In order to reduce this phenomenon, a saturation function is used.

In recent years, great efforts have been made to increase the mechanical robustness and reliability of the induction motor, and to reduce costs and hardware complexity. Thus, it is necessary to eliminate the speed sensor. Several methods of speed estimators have been proposed in the literature among them the Luenberger observer. It is able to provide both rotor speed and flux without problems of closed-loop integration. In this paper, the fuzzy logic controller (FLC) replaces the PI controller in the speed adaption mechanism of the Luenberger observer. The main advantages of the FLC introduced by Zadeh [5] that it does not require accurate mathematical model of the system studied. Fuzzy logic is based on the linguistic rules by means of IF-THEN rules with the human language.

A. Bennassar, A. Abbou, M. Akherraz and M. Barara are with Electrical Engineering Department , Mohamed V University, Mohammadia School's of Engineers, Street Ibn Sina B.P 765 Agdal Rabat Morocco (e-mail: ab.bennassar@gmail.ma, abbou@emi.ac.ma, akherraz@emi.ac.ma, mohamed-barara@hotmail.fr).

II. Induction Motor Oriented Model

In field oriented control, the flux vector is forced to align with d-axis ($\Phi_{rd} = \Phi_r$ and $\Phi_{rq} = 0$). Thus, the dynamic model of the induction motor in (d, q) reference frame can be expressed in the form of the state equations as shown below:

$$\begin{cases} \dfrac{di_{sd}}{dt} = -\gamma i_{sd} + \omega_s i_{sq} + \dfrac{K}{T_r}\varphi_r + \dfrac{1}{\sigma L_s} v_{sd} \\ \dfrac{di_{sq}}{dt} = -\gamma i_{sq} - \omega_s i_{sd} - K\varphi_r \omega_r + \dfrac{1}{\sigma L_s} v_{sq} \\ \dfrac{d\varphi_{rd}}{dt} = \dfrac{L_m}{T_r} i_{sd} - \dfrac{1}{T_r}\varphi_r \\ \dfrac{d\varphi_{rq}}{dt} = \dfrac{L_m}{T_r} i_{sq} - (\omega_s - \omega_r)\varphi_r \\ \dfrac{d\Omega}{dt} = \dfrac{PL_m}{JL_r}\varphi_r i_{sd} - \dfrac{f}{J}\Omega - \dfrac{T_L}{J} \end{cases} \tag{1}$$

Where:

$$\gamma = \frac{R_s}{\sigma L_s} + \frac{1-\sigma}{\sigma T_r}; K = \frac{1-\sigma}{\sigma L_m}; \sigma = 1 - \frac{L_m^2}{L_s L_r}; T_r = \frac{L_r}{R_r}$$

The angular frequency ω_s of the rotor flux is obtained as the sum of the slip frequency ω_{sl} and rotor electrical speed:

$$\omega_s = \omega_r + \omega_{sl} \tag{2}$$

The space angle of the rotor flux is given by:

$$\theta_s = \theta_r + \int \frac{1}{T_r} \frac{i_{sq}}{i_{sd}} \quad (3)$$

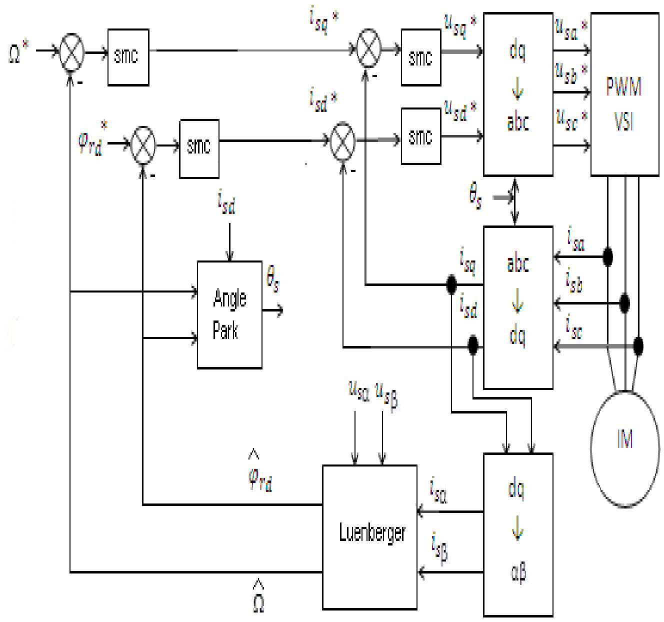


Fig. 1 Block diagram of proposed scheme

III. Sliding Mode Control

Sliding mode technique is a type of variable structure system (VSS) applied to the non-linear systems. The sliding mode control design is to force the system state trajectories to the sliding surface S(x) and to stay on it by means a control defined by the following equation [6]:

$$u = u_{eq} + u_n \quad (4)$$

Where u_{eq} and u_n represent the equivalent control and the discontinue control respectively.

$$u_n = k.sat\left(\frac{s}{\xi}\right) \quad (5)$$

Here ξ defines the thickness of the boundary layer and $sat\left(\frac{s}{\xi}\right)$ is a saturation function.

$$sat\left(\frac{s}{\xi}\right) = \begin{cases} sgn\left(\frac{s}{\xi}\right) si \left|\frac{s}{\xi}\right| > 0 \\ \frac{s}{\xi} si \left|\frac{s}{\xi}\right| < 0 \end{cases} \quad (6)$$

To attract the trajectory of the system towards the sliding surface in a finite time, $u_n(x)$ should be chosen such that Lyapunov function, satisfies the Lyapunov stability:

$$\dot{S}(x)S(x) < 0 \quad (7)$$

The general equation to determine the sliding surface proposed is as follow [7]:

$$S(x) = \left(\frac{d}{dt} + \lambda\right)^{n-1} e \quad (8)$$

Here, e is the tracking error vector, λ is a positive coefficient and n is the system order.

A. Sliding Mode Speed Controller

Considering the equation (8) and taken $n = 1$, the sliding surface of speed can be defined as:

$$S(\Omega) = \Omega^* - \Omega \quad (9)$$

By derivation of equation (9) and taken the fifth equation of the system (1), we obtain:

$$\dot{S}(\Omega) = \dot{\Omega}^* - \frac{PL_m}{JL_r}\varphi_{rd} i_{sq} - \frac{T_L}{J} - \frac{f}{J}\Omega \quad (10)$$

We take:

$$i_{sq} = i_{sq}^{eq} + i_{sq}^{n} \quad (11)$$

During the sliding mode and in permanent regime, $S(\Omega) = \dot{S}(\Omega) = 0$, $i_{sq}^n = 0$. The equivalent control action can be defined as follow:

$$i_{sq}^{eq} = \frac{JL_r}{PL_m\varphi_{rd}}\left(\Omega^* + \frac{T_L}{J} + \frac{f}{J}\Omega\right) \quad (12)$$

During the convergence mode, the condition $\dot{S}(\Omega)S(\Omega) < 0$ must be verified. Therefore, the discontinue control action can be given as:

$$i_{sq}^{n} = k_{isq}.sat\left(\frac{S(\Omega)}{\xi_{isq}}\right) \quad (13)$$

To verify the system stability, coefficient k_{isq} must be strictly positive.

B. Sliding Mode Flux Controller

Considering the equation (8) and taken $n = 1$, the sliding surface of flux can be defined as:

$$S(\varphi_{rd}) = \varphi_{rd}^* - \varphi_{rd} \quad (14)$$

ISBN: 978-1-61804-227-9

By derivation of equation (14) and taken the third equation of the system (1), we obtain:

$$\dot{S}\left(\varphi_{rd}\right)=\dot{\varphi}_{dr}^{*}+\frac{1}{T_r}\varphi_{rd}i_{sq}-\frac{L_m}{T_r}i_{sd} \tag{15}$$

We take:

$$i_{sd}=i_{sd}^{eq}+i_{sd}^{n} \tag{16}$$

During the sliding mode and in permanent regime, $S(\varphi_{rd})=\dot{S}(\varphi_{rd})=0$, $i_{sd}^{n}=0$. The equivalent control action can be defined as follow:

$$i_{sd}^{eq}=\frac{L_r}{L_m}\left(\dot{\varphi}_{dr}^{*}+\frac{1}{T_r}\varphi_{rd}\right) \tag{17}$$

During the convergence mode, the condition $\dot{S}(\varphi_{rd})S(\varphi_{rd})<0$ must be verified. Therefore, the discontinue control action can be given as:

$$i_{sd}^{n}=k_{isd}.sat\left(\frac{S\left(\varphi_{rd}\right)}{\xi_{isd}}\right) \tag{18}$$

To verify the system stability, coefficient k_{isd} must be strictly positive.

C. Sliding Mode Current Controller

Considering the equation (8) and taken $n=1$, the sliding surface of stator currents can be defined as:

$$S\left(i_{sd}\right)=i_{sd}^{*}-i_{sd} \tag{19}$$

$$S\left(i_{sq}\right)=i_{sq}^{*}-i_{sq} \tag{20}$$

By derivation of equation (19) and (20) and taken the first and second equation of the system (1) respectively, we obtain:

$$\dot{S}\left(i_{sd}\right)=\dot{i}_{sd}^{*}+\gamma i_{sd}-\omega_s i_{sq}-\frac{K}{T_r}\varphi_r-\frac{1}{\sigma L_s}v_{sd} \tag{21}$$

$$\dot{S}\left(i_{sq}\right)=\dot{i}_{sq}^{*}+\gamma i_{sq}+\omega_s i_{sd}+K\varphi_r\omega_r-\frac{1}{\sigma L_s}v_{sq} \tag{22}$$

During the sliding mode, $S(i_{sd})=\dot{S}(i_{sd})=0$, $v_{sd}^{n}=0$ and $S(i_{sq})=\dot{S}(i_{sq})=0$, $v_{sq}^{n}=0$. The equivalent control actions can be defined as follow:

$$v_{sd}^{eq}=\sigma L_s\left(\dot{i}_{sd}^{*}+\gamma i_{sd}-\omega_s i_{sq}-\frac{K}{T_r}\varphi_{rd}\right) \tag{23}$$

$$v_{sq}^{eq}=\sigma L_s\left(\dot{i}_{sq}^{*}+\gamma i_{sq}+\omega_s i_{sd}+K\omega_r\varphi_{rd}\right) \tag{24}$$

During the convergence mode, the conditions $\dot{S}(i_{sd})S(i_{sd})<0$ and $\dot{S}(i_{sq})S(i_{sq})<0$ must be verified. Therefore, the discontinue control action can be given as:

$$v_{sd}^{n}=k_{vsd}.sat\left(\frac{S\left(i_{sd}\right)}{\xi_{vsd}}\right) \tag{25}$$

$$v_{sq}^{n}=k_{vsq}.sat\left(\frac{S\left(i_{sq}\right)}{\xi_{vsq}}\right) \tag{26}$$

To verify the system stability, coefficients k_{vsd} and k_{vsq} must be strictly positive.

IV. Luenberger Observer

The Luenberger observer is a deterministic type of observer based on a deterministic model of the system [8]. In this work, the LO state observer is used to estimate the flux components and rotor speed of induction motor by including an adaptive mechanism based on the Lyapunov theory. In general, the equations of the LO can be expressed as follow:

$$\begin{cases}\dot{\hat{x}}=A\hat{x}+Bu+L(y-\hat{y})\\ \hat{y}=C\hat{x}\end{cases} \tag{27}$$

The symbol ^ denotes estimated value and L is the observer gain matrix. The mechanism of adaptation speed is deduced by Lyapunov theory. The estimation error of the stator current and rotor flux, which is the difference between the observer and the model of the motor, is given by [9]:

$$\dot{e}=\left(A-LC\right)e+\Delta A\hat{x} \tag{28}$$

Where:

$$e=x-\hat{x} \tag{29}$$

$$\Delta A=A-\hat{A}=\begin{bmatrix}0&0&0&\mu\Delta\omega_r\\0&0&-\mu\Delta\omega_r&0\\0&0&0&-\Delta\omega_r\\0&0&\Delta\omega_r&0\end{bmatrix} \tag{30}$$

$$\Delta\omega_r=\omega_r-\hat{\omega}_r \tag{31}$$

We consider the following Lyapunov function:

$$V=e^{T}e+\frac{\left(\Delta\omega_r\right)^2}{\lambda} \tag{32}$$

ISBN: 978-1-61804-227-9

Where λ is a positive coefficient. By derivation equation (32), the adaptation law for the estimation of the rotor speed $\hat{\omega}_r$ can be deduced as:

$$\hat{\omega}_r = \int \lambda K \left(e_{is\alpha}\hat{\varphi}_{r\beta} - e_{is\beta}\hat{\varphi}_{r\alpha}\right)dt \qquad (33)$$

The speed is estimated by a PI controller described as:

$$\hat{\omega}_r = K_p\left(e_{is\alpha}\hat{\varphi}_{r\beta} - e_{is\beta}\hat{\varphi}_{r\alpha}\right) + \frac{K_i}{s}\int\left(e_{is\alpha}\hat{\varphi}_{r\beta} - e_{is\beta}\hat{\varphi}_{r\alpha}\right)dt \quad (34)$$

With K_p and K_i are positive constants. The feedback gain matrix L is chosen to ensure the fast and robust dynamic performance of the closed loop observer [10]-[11].

$$L = \begin{bmatrix} l_1 & -l_2 \\ l_2 & l_1 \\ l_3 & -l_4 \\ l_4 & l_3 \end{bmatrix} \qquad (35)$$

With l_1, l_2, l_3 and l_4 are given by:

$$l_1 = (k_1 - 1)\left(\gamma + \frac{1}{T_r}\right)$$
$$l_2 = -(k_1 - 1)\hat{\omega}_r$$
$$l_3 = \frac{(k_1^2 - 1)}{K}\left(\gamma - K\frac{L_m}{T_r}\right) + \frac{(k-1)}{K}\left(\gamma + \frac{1}{T_r}\right)$$
$$l_4 = -\frac{(k-1)}{K}\hat{\omega}_r$$

Where k_1 is a positive coefficient obtained by pole placement approach [12].

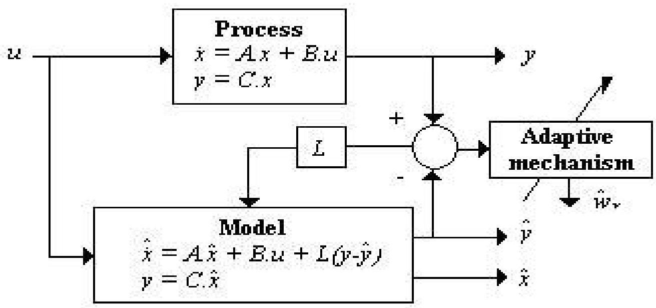


Fig. 2. Block diagram of Luenberger observer

In this paper, we will replace the PI controller in Luenberger observer adaptation mechanism by a fuzzy logic controller.

V. Fuzzy Logic Control

Fig. 3 shows the block diagram of fuzzy logic controller system where the variables K_p, K_i and B are used to tune the controller.

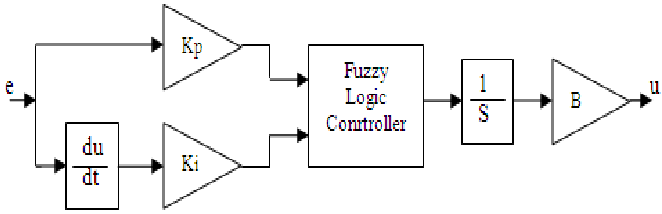


Fig. 3. Block diagram of a fuzzy logic controller

There are two inputs, the error and the change of error. The FLC consists of four major blocks, Fuzzification, knowledge base, inference engine and defuzzification.

A. *Fuzzification*

The crisp input variables are e and ce are transformed into fuzzy variables referred to as linguistic labels. The membership functions associated to each label have been chosen with triangular shapes. The following fuzzy sets are used, NL (Negative Large), NM (Negative Medium), NS (Negative Small), ZE (Zero), PS (Positive Small), PM (positive Medium), and PL (Positive Large). The universe of discourse is set between –1 and 1. The membership functions of these variables are shown in Fig. 4.

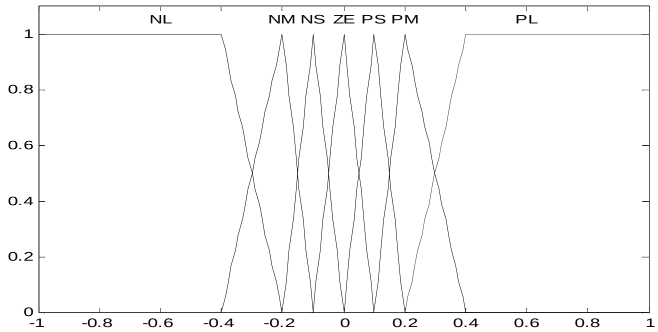


Fig 4. Membership functions

B. *Knowledge Base and Inference Engine*

The knowledge base consists of the data base and the rule base. The data base provides the information which is used to define the linguistic control rules and the fuzzy data in the fuzzy logic controller. The rule base specifies the control goal actions by means of a set of linguistic control rules [16]. The inference engine evaluates the set of IF-THEN and executes 7*7 rules as shown in Table I.

ce/e	NL	NM	NS	ZE	PS	PM	PL
NL	NL	NL	NL	NL	NM	NS	ZE
NM	NL	NL	NL	NM	NS	ZE	PS
NS	NL	NL	NM	NS	ZE	PS	PM
ZE	NL	NM	NS	ZE	PS	PM	PL
PS	NM	NS	ZE	PS	PM	PL	PL
PM	NS	ZE	PS	PM	PL	PL	PL
PL	ZE	PS	PM	PL	PL	PL	PL

Table I. Fuzzy rule base

ISBN: 978-1-61804-227-9

The linguistic rules take the form as in the following example:

IF e is NL AND ce is NL THEN u is NL

C. *Deffuzzification*

In this stage, the fuzzy variables are converted into crisp variables. There are many defuzzification techniques to produce the fuzzy set value for the output fuzzy variable. In this paper, the centre of gravity defuzzification method is adopted here and the inference strategy used in this system is the Mamdani algorithm.

VI. SIMULATION RESULTS AND DISCUSSION

A series of simulation tests were carried out on sliding mode control of induction motor based on the Luenberger observer using fuzzy logic controller in adaptation mechanism. Simulations have been realized under the Matlab/Simulink. The parameters of induction motor used are indicated in Table II.

A. *Operating at Load Torque*

Figures 5, 6 and 7 represent the simulation results obtained from a no load operating. We impose a speed of reference of 100 rad/s and we applied a load torque with 10 N.m between t = 1 s and t = 1.5 s.

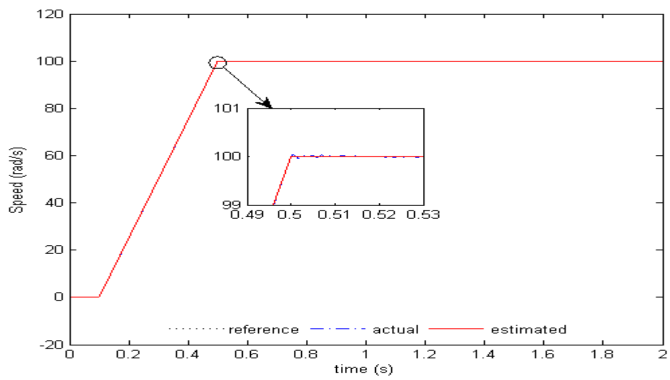


Fig 5. Rotor speed

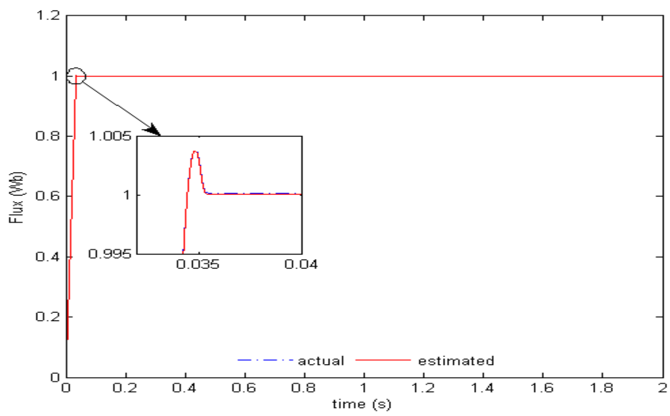


Fig 6. Rotor flux

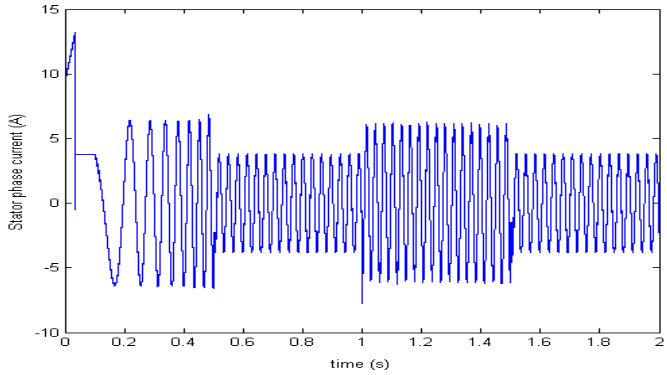


Fig 7. Stator phase current

B. *Operating at Inversion of Speed*

In this case, we applied a speed reference varying between 100 rad/s to -100 rad/s.

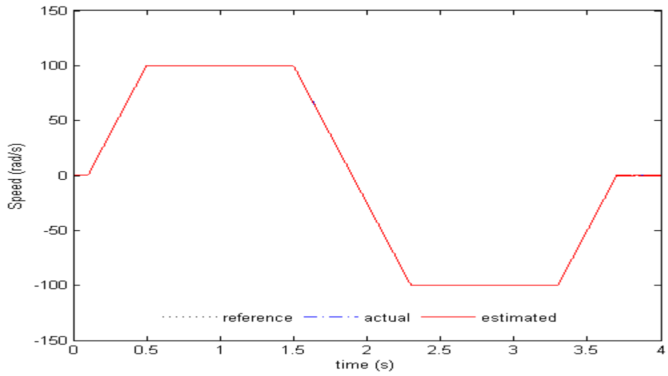


Fig 8. Rotor speed

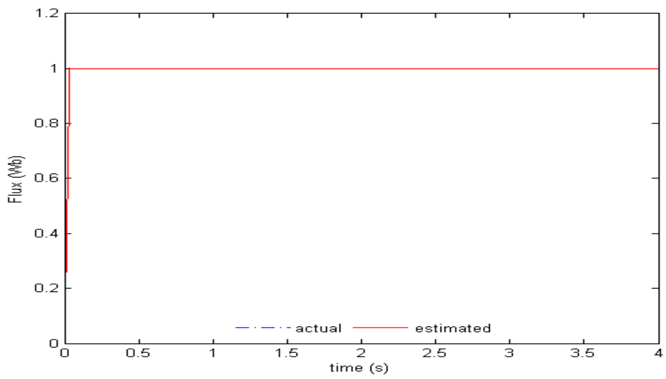


Fig 9. Rotor flux

C. *Operating at Load Speed*

Figures 10 and 11 illustrate simulation results with a speed carried out for low speed ±10 rad/s.

ISBN: 978-1-61804-227-9

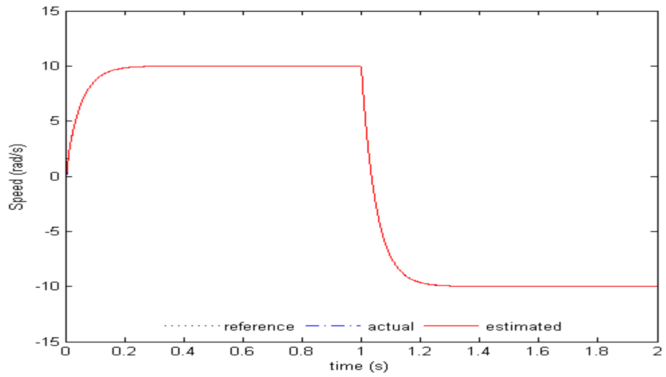


Fig 10. Rotor speed

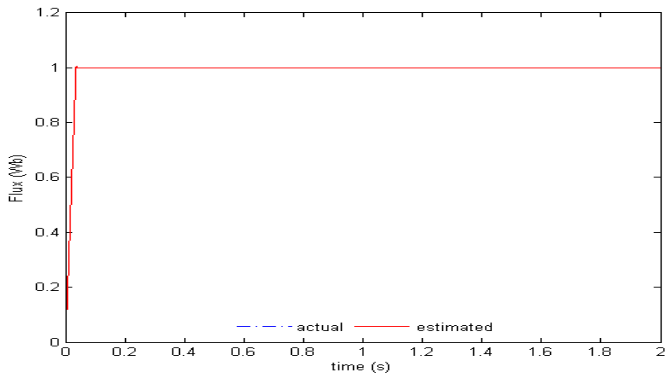


Fig 11. Rotor flux

With the results above, we can see the good estimated speed tracking performance test in different working in inverse and low speed in terms of overshoot, static error and fast response. The flux is very similar to the nominal case. The stator phase current remains sinusoidal and takes appropriate value. It is evident from these simulation results that the proposed sliding mode control presents an excellent performance.

VII. Conclusion

In this paper we have presented the sensorless sliding mode control using the Luenberger observer with fuzzy logic adaptation mechanism. The simulation results have demonstrated the performances of the proposed scheme for steady state responses of flux and speed even at inverse and low speed and with application of the load torque disturbances.

Appendix

Rated power	3 KW
Voltage	380V Y
Frequency	50 Hz
Pair pole	2
Rated speed	1440 rpm
Stator resistance	2.2 Ω
Rotor resistance	2.68 Ω
Inductance stator	0.229 H
Inductance rotor	0.229 H
Mutual inductance	0.217 H
Moment of Inertia	0.047 $kg.m^2$

Table II. Induction motor parameters

References

[1] A. Abbou, T. Nasser, H. Mahmoudi, M. Akherraz, and A. Essadki, “Induction motor control and implementation using dspace,” WSEAS Transactions on Systems and Control, 2012, vol. 7, pp. 26-35.

[2] V.I Utkin, “Variable structure systems with sliding modes,” IEEE Trans. Automat. contr, AC-22, pp. 212–221. February 1993.

[3] A. Ghazel, B. de Fornel, J.C. Hapiot, “Robustesse d’un contrôle vectorielle de structure minimale d’une machine asynchrone,” J. Phys. III France, pp. 943-958, 1996.

[4] C. Vecchio, Sliding Mode Control: Theoretical Developments and Applications to Uncertain Mechanical System, Degli Studi Pavia University.

[5] LA. Zadeh, “Fuzzy sets, Information and Control,” vol. 8, pp. 338-353, 1965.

[6] V. Utkin, “Variable structure systems with sliding modes,” IEEE Transactions on Automatic Control, Vol .22, no. 2, pp. 212–222, 1977.

[7] JJE. Slotine, W. Li, “Applied nonlinear control,” New Jersey, Prentice-Hall, Inc, 1991.

[8] Juraj Gacho and Milan Zalman, “IM based speed servodrive with Luenberger observer,” Journal of Electrical Engineering, vol. 6, no. 3, pp. 149-156, 2010.

[9] J. Maes and J. Melkebeek, “Speed sensorless direct torque control of induction motor using an adaptive flux observer,” Proc. Of IEEE Trans. Industry Appl, vol. 36, pp. 778-785, 2000.

[10] S. Belkacem, F. Naceri, A. Betta and L. Laggoune, “Speed sensorless of induction motor based on an improved adaptive flux observer,” IEEE Trans. Industry Appl, pp. 1192-1197, 2005.

[11] B. Akin, “State estimation techniques for speed sensorless field oriented control of induction motors,” M.Sc. Thesis EE Dept, METU, 2003.

[12] Sio-Iong Ao Len Gelman, “Advances in electrical engineering and computational science lecture,” Notes in Electrical Engineering, vol. 39, Editors, 2009.

[13] A. Bennassar, A. Abbou, M. Akherraz, and M. Barara, “Sensorless backstepping control using an adaptive Luenberger observer with three levels NPC inverter,” WASET, International Journal of Electrical Science and Engineering Vol. 7, no. 8, pp. 1102-1108, 2013.

Abderrahim Bennassar was born in Casablanca, Morocco in 1987. He received Master degree in treatment of information from Hassan 2 University, Casablanca in 2011. Currently, he is pursuing PhD degree at Mohammadia School of Engineering, Rabat.

His researcher interests include the control strategies for AC Drives, especially Induction Motor Drives and Sensorless Control.

Ahmed Abbou received the ‘’ Agrégation Génie Electrique’’ from Ecole Normale Supérieur de l’Enseignement Technique ENSET Rabat in 2000. He received the ‘’Diplôme des Etudes Supérieurs Approfondies’’ in industrial electronics from Mohammadia School’s of engineers, Rabat in 2005. He received with Honors the Ph.D. degree in industrial electronics and electrical machines, from Mohammadia School’s of engineers Rabat in 2009.

Since 2010, he has been a Professor of power Electronic and Electric drives at the Mohammadia School’s of engineers, Rabat. He has presented papers at national and international conference on the Electrical machine, Power Electronic and electric drives. His current area of interest is related to the innovative control strategies for Ac machine Drives, renewable energy.

Mohammed Akherraz Graduated from the Mohammadia School's of engineers Rabat morocco, in 1980. In 1983 he was graduated a Fulbrighr scholarship to pursue his post-graduate studies. He earned the Ph.D degree in 1987 rom UW, Seattle. He joined the EE department Of the Mohammadia School's of engineers, Rabat Morocco, where he's presently a Professor of power electronics and Electric drives.
He published numerous papers in international journal and conferences. His areas of interests are: power electronics, Electric drives, Computer Modeling of power Electronics circuit, and systems drives.

Mohamed Barara was born in Fez, Morocco in 1986. He is received master degree in industrial automated systems engineering from college of sciences Fez, Morocco in July 2011. Currently he is Phd student from Mohammadia School of Engineering.
His researcher interested include, advanced control of wind turbine system, power electronic and renewable energy.

A DFA Approach for Motion Model Selection in Sensor Fusion

Erkan Bostanci

Abstract—This paper investigates the use of different motion models in order to choose the most suitable one, and eventually reduce the Kalman filter errors in sensor fusion for user tracking applications. A Deterministic Finite Automaton (DFA) was employed using the innovation parameters of the filter. Results show that the approach presented here reduces the filter error compared to a static model and prevents filter divergence.

Keywords—Deterministic Finite Automaton (DFA), sensor fusion, GPS, IMU.

I. Introduction

INTEGRATION of data from Global Positioning System (GPS) and Inertial Measurement Unit (IMU) sensors has been well-studied [1] in order to improve upon the robustness of the individual sensors against a number of problems related to accuracy or drift. The Kalman filter (KF) [2] is the most widely-used filter due to its simplicity and computational efficiency [3] especially for real-time user tracking applications.

Attempts have been made to improve the accuracy of the filter using adaptive values for the state and measurement covariance matrices based on the innovation [4] and recently fuzzy logic was used for this task [5], [6]. In some studies [7], [8] used dynamic motion parameters to decide on the dominance of individual sensors for the final estimate.

Alternative approaches suggest using different motion models for recognizing the type of the motion [9]–[13]. Some of these studies (*e.g.* [10], [11] use a Bayesian framework for identifying a scoring scheme for selecting a motion model and some other studies, see [13], apply different motion models concurrently and select one of them according to a probabilistic approach.

This paper presents the selection and use of different motion models according to a DFA model in order to reduce the filter error and ensure faster filter convergence. The rest of the paper is structured as follows: Section II presents the methods used for obtaining positional estimates from individual sensors. The fusion filter which uses these motion estimates in order to produce a single output is presented in Section III. Results are given in Section V and the paper is concluded in Section VI.

II. Finding Position Estimates

Before describing the details of the fusion filter and the DFA approach, it is important to present the calculations used for obtaining individual measurements from the GPS (Phidgets 1040) and IMU sensors (Phidgets Spatial 1056), both low-cost sensors with reasonable accuracy.

E. Bostanci is with Computer Engineering Department, Ankara University, 50. Yil Campus, I Blok, Golbasi, Ankara, 06830, Turkey, e-mail:ebostanci@ankara.edu.tr

A. GPS position estimate

The data obtained from the GPS is in well-known NMEA format and includes position, the number of visible satellites and detailed satellite information for a position P on Earth's surface, as shown in Figure 1.

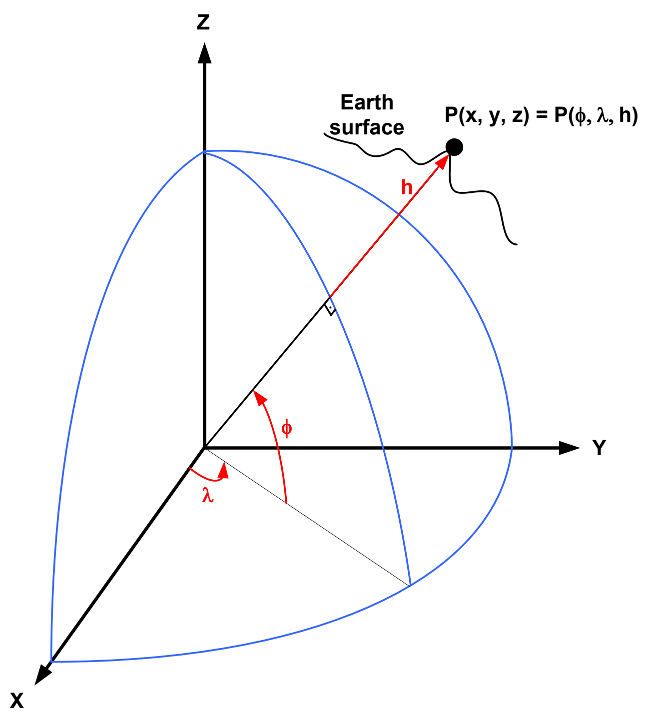


Fig. 1. GPS position parameters in latitude (ϕ), longitude (λ) and altitude (h) and x, y and z in ECEF. Following [14].

Using this information, the GPS coordinates can be converted from geodetic latitude (ϕ), longitude (λ) and altitude (h) notation to ECEF Cartesian coordinates $\mathbf{x_{gps}}$, $\mathbf{y_{gps}}$ and $\mathbf{z_{gps}}$ as:

$$\begin{aligned} x_{gps} &= (N+h)\cos(\phi)\cos(\lambda) \\ y_{gps} &= (N+h)\cos(\phi)\sin(\lambda) \\ z_{gps} &= ((1-e^2)N+h)\sin(\phi) \end{aligned} \tag{1}$$

where

$$N = \frac{a}{\sqrt{1.0 - e^2\sin(\phi)^2}} \tag{2}$$

and a is the WGS84 [15] ellipsoid constant for equatorial earth radius (6,378,137m), e^2 corresponds to the eccentricity of the earth with a value of $6.69437999 \times 10^{-3}$ [3]. The calculated values form the measurements from the GPS sensor as $m_{gps} = (x_{gps}, y_{gps}, z_{gps})$.

B. IMU position estimate

Finding the position estimate from the IMU is performed by double-integrating the accelerometer outputs for several samples, the current implementation uses four samples. The first integration, to find the velocity, involves integrating accelerations using $v(t) = v(0) + at$:

$$\begin{aligned} v_x &= \int_0^T a_x \mathrm{d}t = v_x(T) - v_x(0) \\ v_y &= \int_0^T a_y \mathrm{d}t = v_y(T) - v_y(0) \\ v_z &= \int_0^T a_z \mathrm{d}t = v_z(T) - v_z(0) \end{aligned} \tag{3}$$

Since multiple samples are taken, $\mathrm{d}t$ is the time passed for each one of them. The next step is to integrate the velocities from (3) to find the position using $x(t) = x(0) + vt$ as

$$\begin{aligned} x_{imu} &= \int_0^T v_x \mathrm{d}t = p_x(T) - p_x(0) \\ y_{imu} &= \int_0^T v_y \mathrm{d}t = p_y(T) - p_y(0) \\ z_{imu} &= \int_0^T v_z \mathrm{d}t = p_z(T) - p_z(0) \end{aligned} \tag{4}$$

These calculated positions ($m_{imu} = (x_{imu}, y_{imu}, z_{imu})$) are used as the measurements from the IMU.

III. Fusion Filter

The filter designed for integration of three sensors consists of a state x which includes positional data ($P = (P_x, P_y, P_z)^T$), linear velocities ($V = (V_x, V_y, V_z)^T$):

$$x = (P, V)^T \tag{5}$$

A simple state consisting of 6 elements will facilitate obtaining a better performance in speed than one with a larger state. At each iteration, the predict–measure–update cycle of the KF is executed in order to produce a single output from several sensors as the filter output.

In the first stage, *i.e.* prediction, a transition matrix (F of (6)) is applied to the state x in order to obtain the predicted position:

$$F = \begin{bmatrix} 1 & 0 & 0 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 1 & 0 & 0 & \Delta t \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{6}$$

where Δt is the time between two prediction stages.

Measurements are obtained from the GPS and the IMU using the values obtained as described in Section II and are combined to create a measurement vector:

$$z = (x_{gps} + x_{imu}, y_{gps} + y_{imu}, z_{gps} + z_{imu})^T \tag{7}$$

Here, the IMU measurements for position are used as offsets to the position obtained from the most recent GPS fix.

IV. DFA based Model Transitions

The difference between the measurements (z) and the prediction ($h\hat{x}$), omitting the subscripts indicating time, is defined as the innovation (y):

$$y = z - h\hat{x} \tag{8}$$

The innovation vector has 3 components for position elements as y_x, y_y and y_z. The DFA model presented here uses the magnitude of these to define the filter divergence as

$$I = \sqrt{y_x^2 + y_y^2 + y_z^2} \tag{9}$$

and uses the following rules to assign the values of I into different classes named I_0, I_1 and I_2 which are defined as

$$I = \begin{cases} I_0 : & I < 3.0 \\ I_1 : & 3.0 \leq I < 7.5 \\ I_2 : & 7.5 \leq I \end{cases} \tag{10}$$

A DFA consists of several elements which can be listed as states, input symbols and transition rules [16]. The states of the DFA defined in Figure 2 correspond to different motion models.

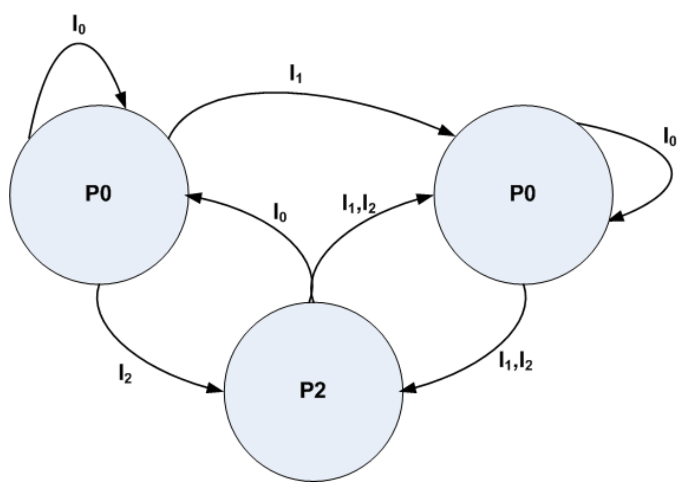


Fig. 2. DFA model for the model transitions

The classes (I_0, I_1 and I_2) are considered as the input symbols used for the DFA model. Finally, the transitions between states model the selection mechanism presented in this paper.

When a model (P_i) is selected, the value for i is used as a velocity coefficient (c_i) in the transition function (F) for position ($\hat{x}_P = x_P + c_i V \Delta t$). For instance, `P0` indicates a stationary transition model where the current values for position (`P`) will be unchanged in the predicted state, whereas `P2` indicates a motion model where position is predicted with twice the current positional velocities ($\hat{x}_P = x_P + (2 \times V)\Delta t$) in order to adapt any sudden changes in the estimated position.

During experiments it was observed that in some cases selected models could be changing very often. A sliding window filter was applied to the results of the model selection logic in order to prevent frequent transitions between different motion models. In the implementation, the most recent five models were averaged to obtain the final motion model as illustrated in Figure 3.

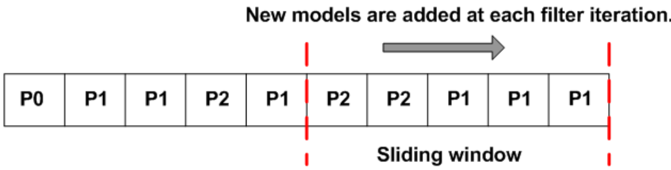


Fig. 3. Sliding window for preventing frequent model transitions

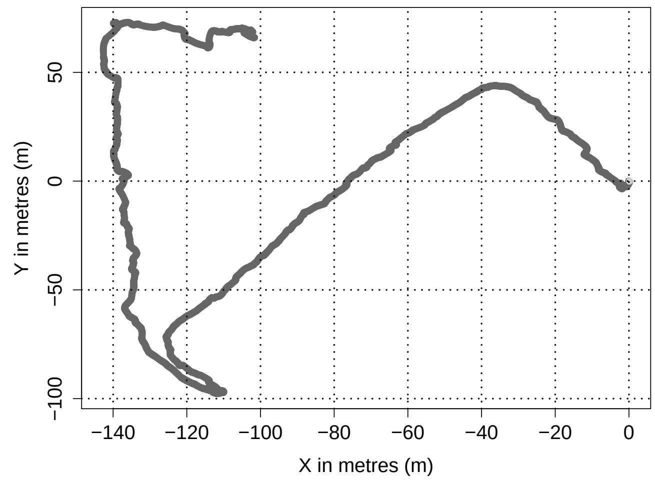
V. Results

Experiments were conducted using low-cost GPS and IMU sensors mounted on a cycle helmet for a user walking with varying speed. Sampling rate for the IMU was selected as 20 milliseconds and a GPS fix was received every second.

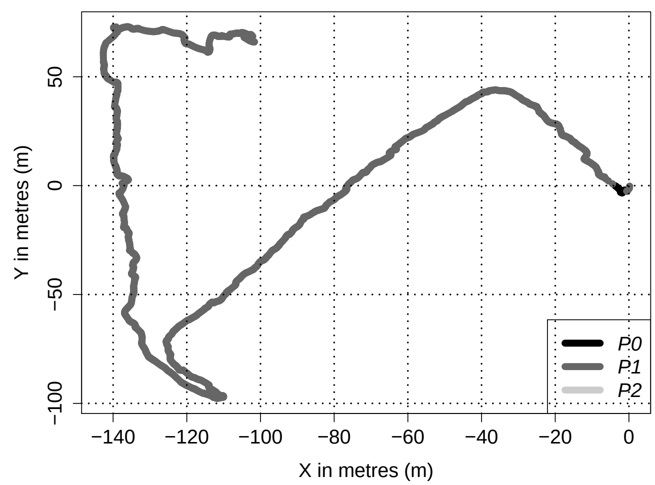
Figures 4 to 6 show the estimated paths using integration of the two sensors and employing different motion models. Portions of the estimated paths are coloured differently in order to indicate the type of the motion model used for estimation. It is important to note that the static model used in the results correspond to `P1` and hence is drawn in the same colour.

Data 2 was acquired while the sensors were completely stationary, the accuracy of the sensor fusion is found in this case as 1.5m which is, indeed, less than the accuracy of the GPS used in the experiments (given as 2.5m in the product specification) — a benefit of sensor fusion. The DFA model selection logic reduced this error even further since the motion model is recognized as P0 (see Figure 5).

Filter errors are presented in Figures 7 to 9. Note that these errors are an indicator of the difference between the filter predictions and the actual values of the measurements. It can be seen that the filter error is reduced when the DFA models are employed.



(a)



(b)

Fig. 4. Trajectory results for Dataset 1. (a) Static motion model (b) DFA models

VI. Conclusion

This paper presented a DFA design for motion model selection in GPS–IMU sensor fusion. The results show that multiple-motion model sensor fusion can be achieved by utilising Kalman filter innovation together with DFA based model selection scheme. It was observed that the use of different motion models can reduce the filter error and prevent divergence. It is clear that selection of the appropriate motion model depending on user's speed improves the accuracy of Kalman filter for tracking applications.

Future work will delve into further analysis of different motion models and a machine learning approach appears to be a promising research direction.

References

[1] P. Groves, *Principles of GNSS, inertial, and multi-sensor integrated navigation systems*, ser. GNSS technology and applications series. Artech House, 2008.

[2] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME Journal of Basic Engineering*, no. 82 (Series D), pp. 35–45, 1960.

[3] E. Kaplan and C. Hegarty, *Understanding GPS: Principles and Applications*, ser. Artech House mobile communications series. Artech House, 2005.

[4] A. Almagbile, J. Wang, and W. Ding, "Evaluating the performances of adaptive kalman filter methods in GPS/INS integration," *Journal of Global Positioning Systems*, vol. 9, no. 1, pp. 33–40, 2010.

[5] C. Tseng, C. Chang, and D. Jwo, "Fuzzy adaptive interacting multiple model nonlinear filter for integrated navigation sensor fusion," *Sensors*, vol. 11, pp. 2090–2111, 2011.

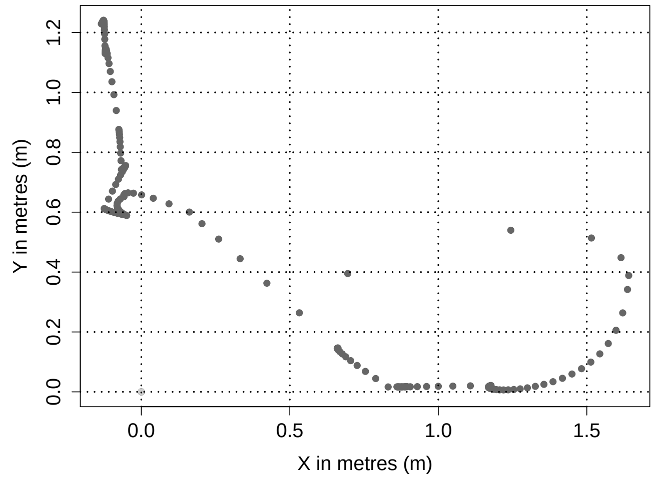
[6] J. Kramer and A. Kandel, "On accurate localization and uncertain sensors," *International Journal of Intelligent Systems*, vol. 27, no. 5, pp. 429–456, 2012.

[7] L. Ojeda and J. Borenstein, "Flexnav: fuzzy logic expert rule-based position estimation for mobile robots on rugged terrain," in *Robotics and Automation, 2002. Proceedings. ICRA '02. IEEE International Conference on*, vol. 1, 2002, pp. 317–322 vol.1.

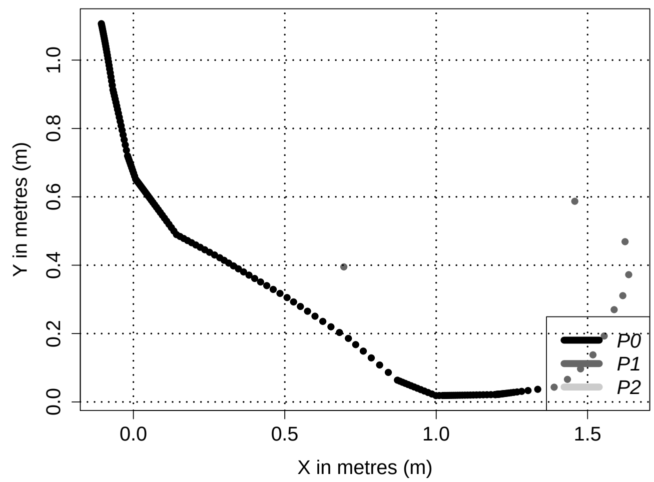
[8] S. K. Hong, "Fuzzy logic based closed-loop strapdown attitude system for unmanned aerial vehicle (uav)," *Sensors and Actuators A: Physical*, vol. 107, no. 2, pp. 109 – 118, 2003.

[9] J. Chen and A. Pinz, "Structure and motion by fusion of inertial and vision-based tracking," in *Austrian Association for Pattern Recognition*, vol. 179, 2004, pp. 75–62.

[10] P. Torr, "Bayesian model estimation and selection for epipolar geometry and generic manifold fitting," *International Journal of Computer Vision*, vol. 50, no. 1, pp. 35–61, 2002.

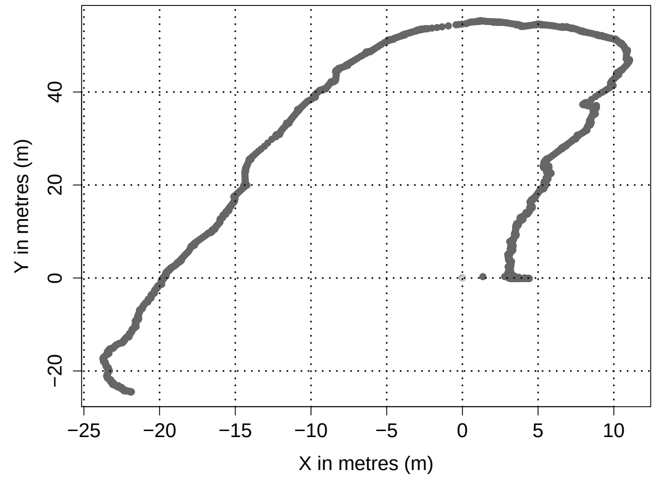


(a)

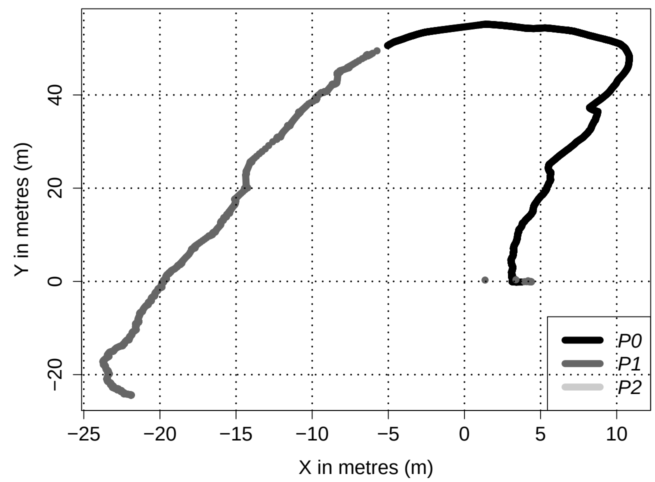


(b)

Fig. 5. Trajectory results for Dataset 2. (a) Static motion model (b) DFA models



(a)



(b)

Fig. 6. Trajectory results for Dataset 3. (a) Static motion model (b) DFA models

[11] K. Kanatani, "Uncertainty modeling and model selection for geometric inference," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 10, pp. 1307–1319, 2004.

[12] K. Schindler and D. Suter, "Two-view multibody structure-and-motion with outliers through model selection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 6, pp. 983–995, 2006.

[13] J. Civera, A. Davison, and J. M. M. Montiel, "Interacting multiple model monocular slam," in *International Conference on Robotics and Automation*, 2008, pp. 3704–3709.

[14] S. H. Stovall, "Basic inertial navigation," Naval Air Warfare Center Weapons Division, Technical Report, 2008.

[15] National Imagery and Mapping Agency, "Department of Defense World Geodetic System 1984: its definition and relationships with local geodetic systems," National Imagery and Mapping Agency, Tech. Rep., 2000.

[16] J. Hopcroft, R. Motwani, and J. Ullman, *Introduction to automata theory, languages, and computation*. Pearson/Addison Wesley, 2007. [Online]. Available: http://books.google.com.tr/books?id=6ytHAQAAIAAJ

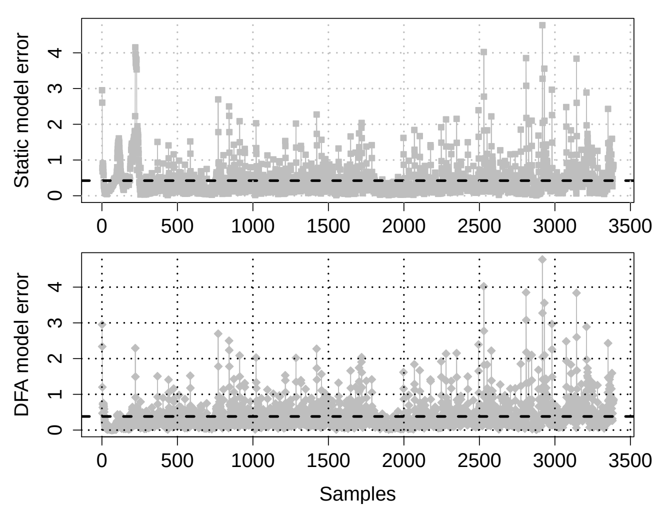


Fig. 7. Filter errors for Dataset 1

ISBN: 978-1-61804-227-9

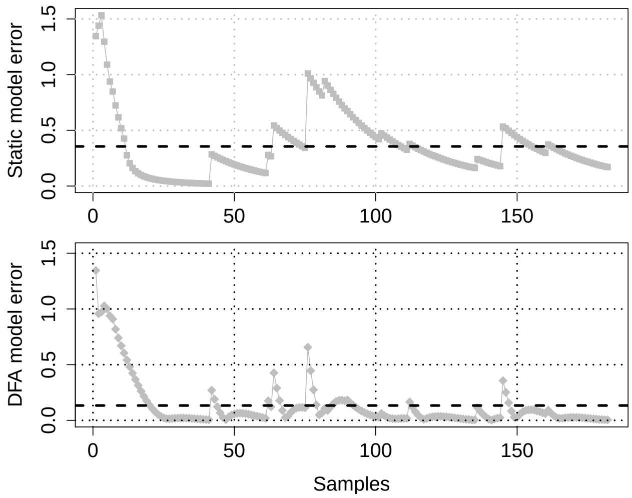


Fig. 8. Filter errors for Dataset 2

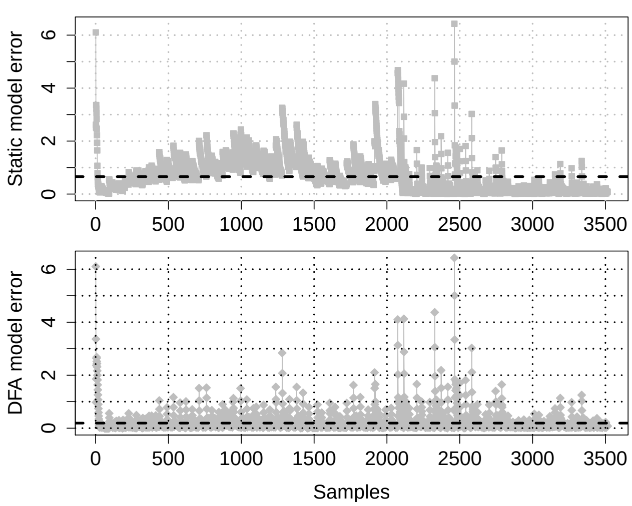


Fig. 9. Filter errors for Dataset 3

ISBN: 978-1-61804-227-9

Fuzzy Logic and Artificial Neural Network associated with DTFC strategy for controlling an induction motor

A. Abbou, A. Bennassar, H. Mahmoudi, and M. Akherraz

***Abstract*—**We propose a Sensorless Direct Torque and Flux Control (DTFC) of Induction Motor (IM) using two approach intelligent techniques: Fuzzy Logic is used for controlling the rotor speed and Artificial Neural Network (ANN) applied in switching select stator voltage. We estimated the rotor speed by using the Model Reference Adaptive Systems (MRAS). The control method proposed in this paper can reduce the torque, stator flux and current ripples and especially improve system good dynamic performance and robustness in high and low speeds.

***Keywords*—** Fuzzy Logic, Artificial Neural Network, Induction Motor, Sensorless DTFC, Controller, MRAS.

I. Introduction

THE induction motor is one of the most widely used machines in industrial applications due to its high reliability, relatively low cost, and modest maintenance requirements. High performance electric drives require decoupled torque and flux control. This control is commonly provided through Field Oriented Control (FOC), which is based on decoupling of the torque-producing current component and the flux- producing component. FOC drive scheme requires current controllers and coordinate transformations. Current-regulated pulse-width-modulation inverter and inner current loops degrade the dynamic performance in the operating regimes wherein the voltage margin is insufficient for the current control, particularly in the field weakening region [1].

The problem of decoupling the stator current in a dynamic fashion is avoided by DTFC. Direct Torque and Flux Control (DTFC) is nowadays widely used for induction motor drives, her provides a very quick and precise torque response without the complex field orientation block and the inner current regulation loop [2]. The disadvantages of conventional DTC are high torque ripple and slow transient response to the step changes in torque during start-up [3]. For that reason the application of Fuzzy logic and artificial neural network attracts the attention of many scientists from all over the word [4]. The reason for this trend is the many advantages which the architectures of ANN have over traditional algorithmic methods [5]. Among the advantages of ANN are the ease of training and generalization, simple architecture, possibility of approximating non linear functions, insensitivity to the distortion of the network, and inexact input data [4, 6]. On the other hand, ongoing research has concentrated on the elimination of the speed sensor at the machine shaft without deteriorating the dynamic performance of the drive control system [7]. The advantages of speed sensorless induction motor drives are reduced hardware complexity and lower cost, reduced size of the drive machine, elimination of the sensor cable, better noise immunity, increased reliability and less maintenance requirements.

In this paper we present the performance of the sensorless speed control of induction motor using a speed proportional integral (PI) Fuzzy controller. The artificial neural network then replaces the switching table of the conventional DTFC while the rotation speed is estimated by the MRAS method. This paper organized as follows: The induction model is presented in the second section, the DTFC based Artificial Neural Network is developed in the third section, the speed PI Fuzzy controller design is performed in the fourth section, section five present a speed MRAS estimator and section six is devoted to illustrating by simulation the performances of this control strategy, a conclusion and reference list end the paper.

II. Induction Motor Model

The state equation of induction motor written in stator reference frame, (α, β) coordinates, can be expressed as follows:

$$\begin{cases} \dot{X} = A(\omega).X + B.U \\ Y = C.X \end{cases} \qquad (1)$$

Where A, B and C are the evolution, the control and the observation matrices respectively.

$$X = [i_{s\alpha} \quad i_{s\beta} \quad \Phi_{s\alpha} \quad \Phi_{s\beta}] \; ; U = \begin{bmatrix} V_{s\alpha} \\ V_{s\beta} \end{bmatrix} ; Y = \begin{bmatrix} i_{s\alpha} \\ i_{s\beta} \end{bmatrix}$$

$$A = \begin{bmatrix} -(\frac{1}{\sigma T_s} + \frac{1-\sigma}{\sigma T_r}) & 0 & \frac{1-\sigma}{\sigma M T_r} & \frac{1-\sigma}{\sigma M}\omega \\ 0 & -(\frac{1}{\sigma T_s} + \frac{1-\sigma}{\sigma T_r}) & -\frac{1-\sigma}{\sigma M}\omega & \frac{1-\sigma}{\sigma M T_r}\omega \\ \frac{M}{T_r} & 0 & -\frac{1}{T_r} & -\omega \\ 0 & \frac{M}{T_r} & \omega & -\frac{1}{T_r} \end{bmatrix}$$

ISBN: 978-1-61804-227-9

$$B=\begin{bmatrix}\frac{1}{\sigma L_s} & 0\\ 0 & \frac{1}{\sigma L_s}\\ 0 & 0\\ 0 & 0\end{bmatrix}\qquad C=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\end{bmatrix}$$

With, ω Rotor speed and the machine's parameters: R_s, R_r are respectively the stator and the rotor resistance, M, L_s, L_r are respectively the mutual, the stator and the rotor cyclic inductances; p denotes the number of pole pairs, with:

$$T_r=\frac{L_r}{R_r},\ T_s=\frac{L_s}{R_s},\ \sigma=1-\frac{M^2}{L_sL_r}$$

The mechanical equation is the following:

$$J.\frac{d}{dt}\Omega=C_e-C_r-f.\Omega \tag{2}$$

In which J is the inertia coefficient and C_r is the load torque. Using the Laplace transform, the equation (3) shows that the relation between the stator flux and the rotor flux represents a low pass with time constant σT_r.

$$\overline{\Phi}_r=\frac{M}{L_s}\frac{\overline{\Phi}_s}{1+\sigma T_r s} \tag{3}$$

The electromagnetic torque can be expressed as

$$C_e=\frac{3}{2}\frac{p}{2}(\Phi_{s\alpha}.i_{s\beta}-\Phi_{s\beta}.i_{s\alpha}) \tag{4}$$

III. DTFC STRATEGY FOR INDUCTION MOTOR

A. *DTFC strategy*

The block diagram of the proposed sensorless control scheme is shown in figure 1.The DTFC method was introduced in 1985 by Takahashi [8]. This strategy of control is relatively new and competitive compare to the rotor flux oriented method.

This type of control is based on the directly determination of the sequence of control applied to the switches of a tension inverter. This choice is generally based on the use of hysteresis regulators, whose function is to control the state of the system, and to modify the amplitude of the stator flux and the electromagnetic torque.

The stator flux, as given in equation (5), can be approximated as equation (6) over a short time period if the stator resistance is ignored.

$$\overline{\Phi}_s=\overline{\Phi}_{so}+\int_0^t(\overline{V}_s-R_s\overline{I}_s)dt \tag{5}$$

$$\overline{\Phi}_s\approx\overline{\Phi}_{so}+\int_0^t\overline{V}_s dt \tag{6}$$

During one period of sampling Te, vector tension applied to the machine remains constant, and thus one can write

$$\overline{\Phi}_s(k+1)\approx\overline{\Phi}_s(k)+\overline{V}_s.T_e \tag{7}$$

$$\text{Or}\quad \Delta\overline{\Phi}_s\approx\overline{V}_s.T_e \tag{8}$$

Therefore to increase the stator flux, we can apply a vector of tension that is co-linear in its direction and vice-versa.

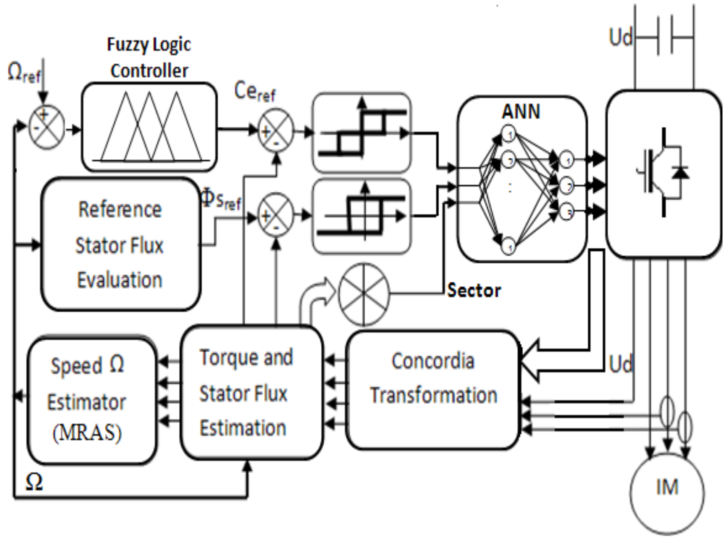


Fig.1.Block diagram of sensorless control proposed

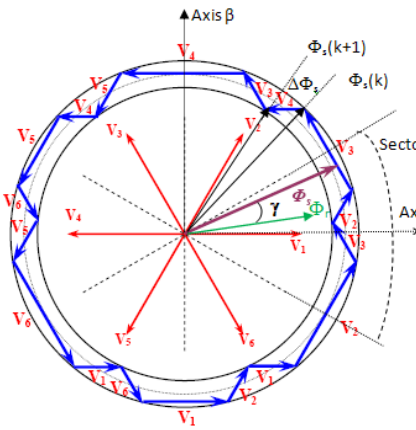


Fig. 2. Definition of stator flux increment

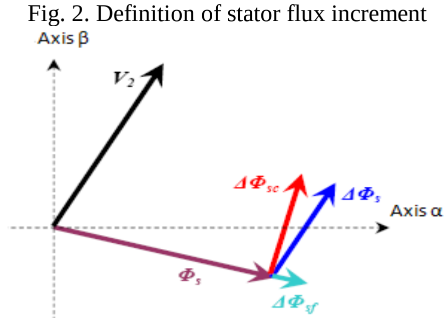


Fig. 3. Components of the error of flux at the time of the application of the vector V_2 voltage

If the error of flux is projected on the direction of stator flux and on a perpendicular direction (Fig.3), one puts in evidence the components acting on the torque and on the flux. In the Figure 3, the component $\Delta\Phi_{sc}$ gives the electromagnetic Torque of the Induction motor while the component $\Delta\Phi_{sf}$ modifies the magnitude of stator flux.

The torque is produced by the induction motor can be expressed as equation:

$$Ce=\frac{3}{2}p\frac{M}{\sigma L_sL_r}\Phi_s\Phi_r\sin\gamma \tag{9}$$

The torque depends upon the amplitude of the two vectors stator flux $\overline{\Phi}_s$ and rotor flux $\overline{\Phi}_r$, and their relative position γ.

ISBN: 978-1-61804-227-9

If one succeeds in perfectly controlling the flux $\overline{\Phi}_s$ (starting with Vs) in module and in position, one can subsequently control the amplitude and the relative position of Φ_r and therefore ultimately, the torque.

When flux is in sector S_i, the vectors V_{i+1or} V_{i-1} are selected to increase the amplitude of flux, and V_{i+2} or V_{i-2} to decrease it. What shows that the choice of the vector tension depends on the sign of the error of flux, independently of its amplitude. This explains why the exit of the corrector of flux can be a Boolean variable. One adds a bond of hysteresis around zero to avoid useless commutations when the error of flux is small. Indeed, with this type of corrector in spite of his simplicity, one can easily control and maintain the end of the vector flux, in a circular ring. The switching table proposed by Takahashi [8], as given by Table 1.

$\Delta\phi_s$	ΔC_e	S_1	S_2	S_3	S_4	S_5	S_6
1	1	110	010	011	001	101	100
	0	000	000	000	000	000	000
	-1	101	100	110	010	011	001
0	1	010	011	001	101	100	110
	0	000	000	000	000	000	000
	-1	001	101	100	110	010	011

Table1. Switching table

B. *Artificial Neural Network DTFC*

The Artificial neural network replaces the switching table selector bock. He uses a dense interconnection of computing nodes to approximate nonlinear function [6]. The neural network selector inputs proposed are the position of flux stator vector represented by the number of the corresponding sector, the error between its estimated value and the reference value and the difference between the estimated electromagnetic torque and the torque reference that is to say three neurons of the input layer (Figure 4). The output layer is composed of three neurons, each representing the state E_i of one of the three pairs of switches T_i of the inverter connected to the positive DC bus.

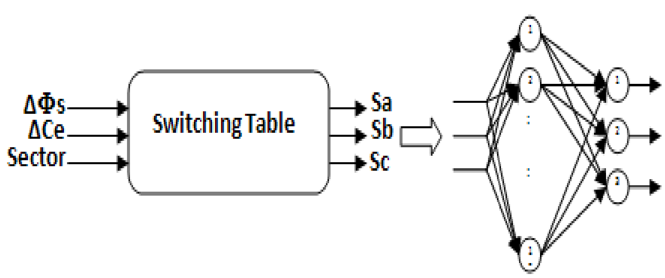


Fig.4. ANN switching Table

After several tests we take an architecture 3-12-3 with a single hidden layer.
The function f (n) activation of the hidden layer is Tansig:

$$a = \frac{e^{n} - e^{-n}}{e^{n} + e^{-n}} \qquad (10)$$

While the activation function of the output layer is Purelin.

$$a = n \qquad (11)$$

The learning of the neural network is done by using the algorithm LVM (levenberg Marquardt) with a number of epochs 500 and an error of 10^{-3}.

IV. Fuzzy logic controller

The block diagram of the PI Fuzzy controller is shown in fig. 5, where the variables K_p, K_i and B are used to tune the controller.

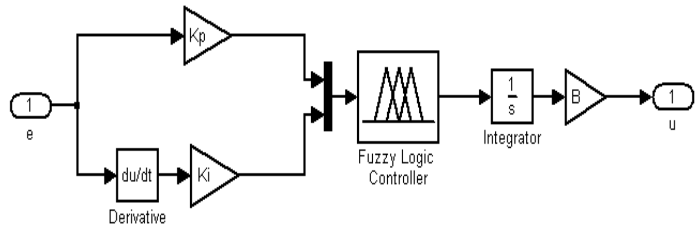


Fig.5. PI Fuzzy controller

One possible initial rule base, that can be used in drive systems for a fuzzy logic controller, consist of 49 linguistic rules, as shown in Table 2, and gives the change of the output of fuzzy logic controller in terms of two inputs: the error (e) and change of error (de). The membership functions of these variables are given in Fig.6:

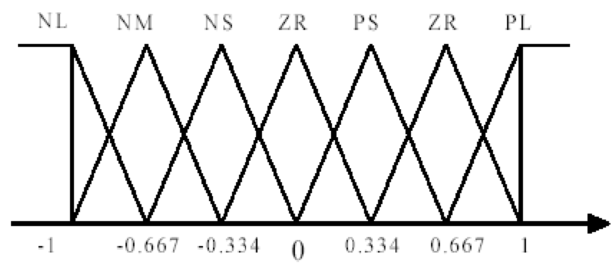


Fig. 6. Membership functions

In Table 2, the following fuzzy sets are used: NL negative large, NM negative medium, NS negative small, ZR zero, PS positive small, PM positive medium and PL positive large. For example, it follows from Table 2 that the first rule is:

IF **e** is NL and **de** is NL then **du** is NL

E/dE	NL	NM	NS	ZR	PS	PM	PL
PL	ZR	PS	PM	PL	PL	PL	PL
PM	NS	ZR	PS	PM	PL	PL	PL
PS	NM	NS	ZR	PS	PM	PL	PL
ZR	NL	NM	NS	ZR	PS	PM	PL
NS	NL	NL	NM	NS	ZR	PS	PM
NM	NL	NL	NL	NM	NS	ZR	PS
NL	NL	NL	NL	NL	NM	NS	ZR

Table 2.Fuzzy rules base

The linguistic rules are in the form of IF-THEN rules and take form: IF (e is X and de is Y) then (du is Z), where X, Y, Z are fuzzy subsets for the universe of discourse of the error, change of error and change of the output. For example, X can denote the subset NEGATIVE LARGE of the error etc. On every of these universes is placed seven triangular membership functions (fig.6). It was chosen to set these universes to normalized type for all of inputs and output. The range of universe is set to -1 to 1.

V. SPEED ESTIMATION WITH MRAS

A linear state observer for the rotor flux can then be derived as follows by considering the mechanical speed as a constant parameter since its variation is very slow in comparison with the electrical variables.

The symbol ˆ denotes an estimated quantity.

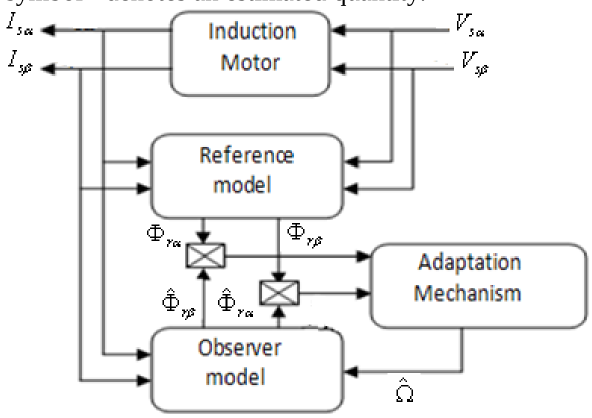


Fig.7. Adaptive observer structure

Since the reference model doesn't depend on the rotation speed, it allows to calculate the components of rotor flux from the equations of the stator voltage:

$$\begin{aligned}\frac{d\Phi_{r\alpha}}{dt} &= \frac{1}{a}(V_{s\alpha} - R_s I_{s\alpha} - \frac{1}{\delta}\frac{dI_{s\alpha}}{dt}) \\ \frac{d\Phi_{r\beta}}{dt} &= \frac{1}{a}(V_{s\beta} - R_s I_{s\beta} - \frac{1}{\delta}\frac{dI_{s\beta}}{dt})\end{aligned} \quad (12)$$

With $\delta = \frac{1}{\sigma L_s}$, $a = \frac{M}{L_r}$ $\quad k = \frac{R_r}{L_r}$

The observer model uses the speed of rotation in its equations and permits to estimate the components of rotor flux:

$$\begin{aligned}\frac{d\hat{\Phi}_{r\alpha}}{dt} &= -k.\Phi_{r\alpha} - p\hat{\Omega}.\Phi_{r\beta} + k.M.I_{s\alpha} \\ \frac{d\hat{\Phi}_{r\beta}}{dt} &= -k.\Phi_{r\beta} + p\hat{\Omega}.\Phi_{r\alpha} + k.M.I_{s\beta}\end{aligned} \quad (13)$$

The adaptation mechanism compares the two models and estimates the speed of rotation by an integral proportional regulator. Using Lyapounov stability theory, we can construct a mechanism to adapt the mechanical speed from the asymptotic convergence's condition of the state variables estimation errors.

$$\hat{\Omega} = K_p.(\hat{\Phi}_{r\alpha}\Phi_{r\beta} - \Phi_{r\alpha}\hat{\Phi}_{r\beta}) + K_i\int(\hat{\Phi}_{r\alpha}\Phi_{r\beta} - \Phi_{r\alpha}\hat{\Phi}_{r\beta})dt \quad (14)$$

K_p and K_i are positive gains.

VI. SIMULATION RESULTS

A. *Speed control*

To study the performance of the sensorless speed control, PI fuzzy controller of speed and neural network switching table with direct torque control strategy, the simulation of the system was conducted using Matlab/Simulink , Fuzzy logic Toolbox and neural network Toolbox. The torque and flux hysteresis bands of 0.5Nm and 0.01Wb respectively were used to give a switching frequency close to 10 kHz at the chosen motor speed and load. The stator flux reference is fixed to 1Wb and we impose a speed of reference varying between -1146 rpm and +1146 rpm:

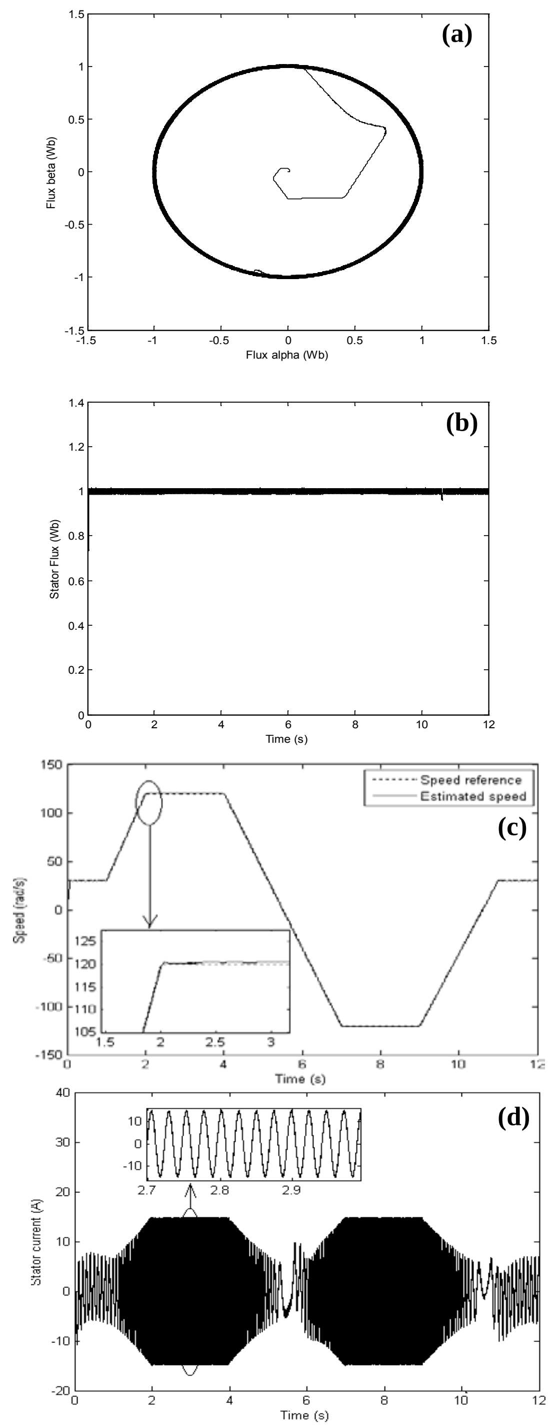


Fig.8. (a) Trajectory of stator flux vector (b) Stator flux magnitude (c) Rotor speed (d) Stator current

B. *Low speed functioning*

We set the rotor flux to 1Wb and verify the operation at low speed between 10rad/s and -10rad/s;

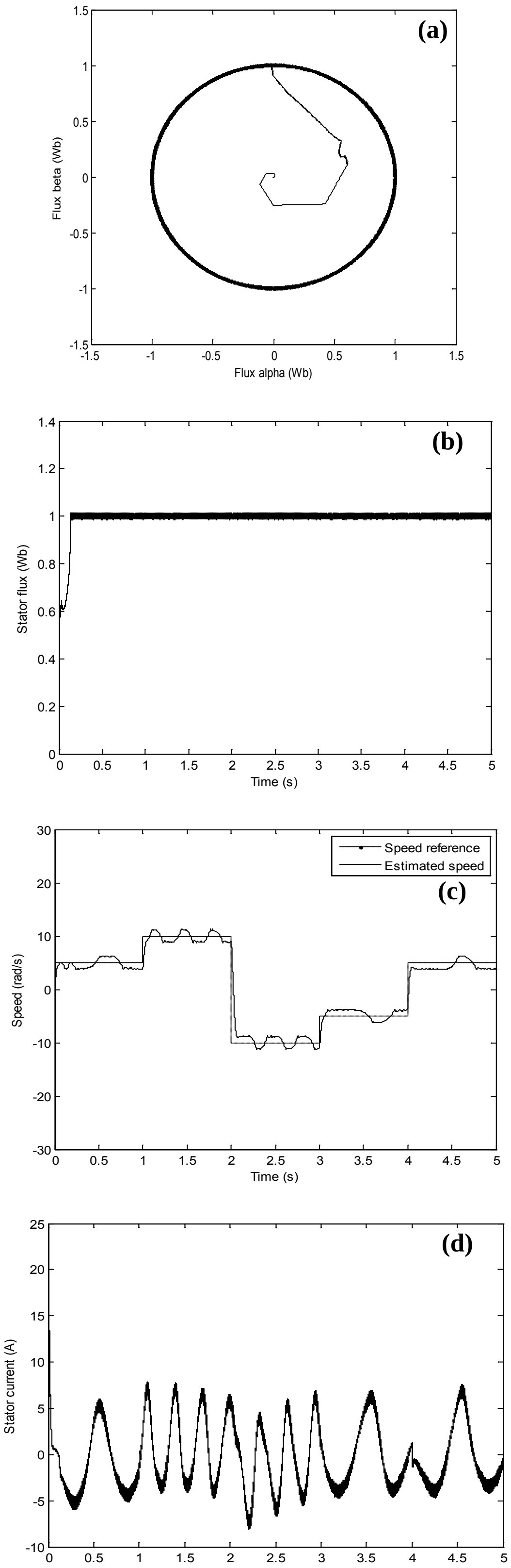


Fig.9. (a) Trajectory of stator flux vector (b) Stator flux magnitude (c) Rotor speed (d) Stator current

C. *High speed functioning*

In the high-speed, we use the law of change of stator flux depending on the rotor speed shown in the following figure [8, 11].

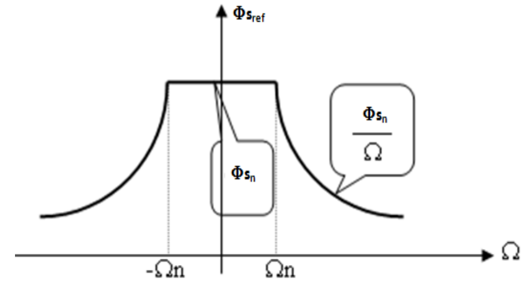


Fig.10: Reference rotor flux law

With Ωn: nominal speed of rotation,

Φ_{sn}: face value of the stator flux.

We keep the same condition of the speed control and impose a rotor speed of 1720 rpm:

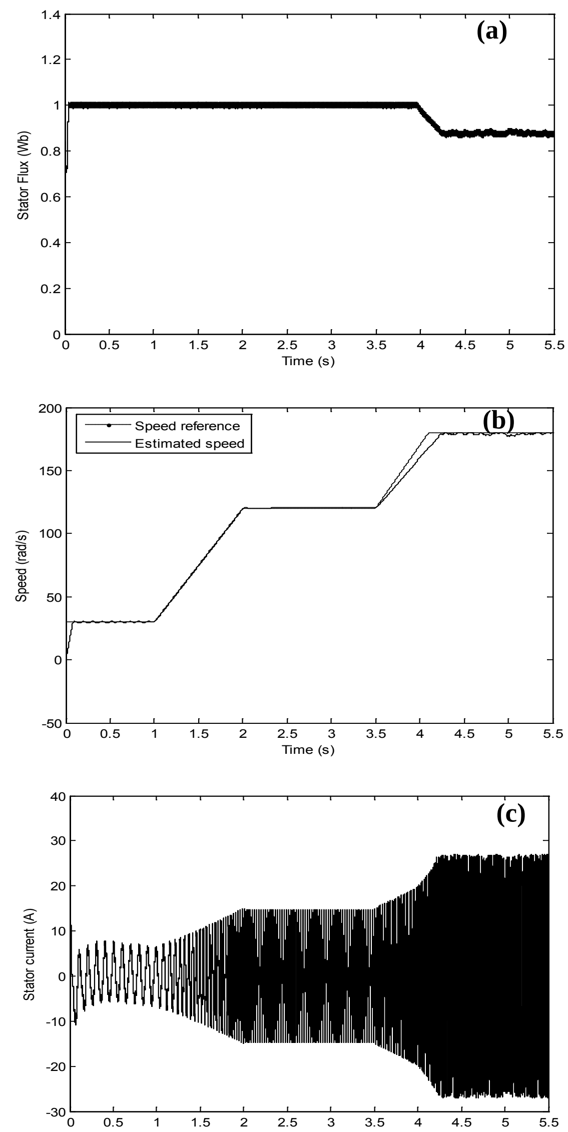


Fig.11. (a) Stator flux magnitude (b) Rotor speed (c) Stator current

ISBN: 978-1-61804-227-9

Figure 8 shows the estimated speed, stator flux and stator current with DTFC scheme and MRAS technique. Estimated speed follows the reference speed closely. The stator phase current in the induction motor remains sinusoidal and takes appropriate value.

Figure 9 shows simulation results of the proposed approach at low speed (±10rad/s). We can see good dynamic behavior and steady state responses of flux and speed. Some estimated speed oscillations can be observed.

Figure 11 illustrate the performance of this strategy in high speed functioning. The estimated speed follows perfectly the speed reference. It is important to note that the control system remain stable.

With the obtained results we can estimate the rotor speed in the different working especially in low and high speed**.** The dynamics of stator flux and the speed estimated are better. The stator phase current in the induction motor remains sinusoidal and takes appropriate value. This strategy of control reduces remarkably the flux, current and torque ripples. The stator flux vector describes a trajectory almost circular.

VII. Conclusion

The control strategy that we have introduced in this paper presents the following advantages:

- Operating without speed sensor.
- good dynamic behavior and steady state responses of speed and flux even at low and high speed
- The stability of system.
- Limitation of the current amplitude and low distortions for current and torque.
- No flux droppings caused by sector changes circular trajectory.
- Flux and torque ripples reduced.

In comparison with some strategies of control presented in literature (Field oriented control, basic DTC, DTC without fuzzy logic controller), this strategy makes the induction motor based DTFC more robust, more stable and good dynamic performance even in high speed.

Appendix

Rated speed N	1440 rpm
Rated voltage V	220 V
Stator resistance Rs	0.85 Ω
Rotor resistance Rr	0.16 Ω
Stator inductance Ls	0.160 H
Rotor inductance Lr	0.023 H
Mutual inductance M	0.058 H
Motor-load inertia J	0.050 $kg.m^2$
Number of pole pairs p	2

Table 3. Induction motor data

References

[1] D.Casadei, F. Profumo, G. Serra and A.Tani, FOC and DTC:Tox viable schemes for induction motors Torque control, IEEE Trans. On Power electronics, Vol.17, No.5, September 2002.

[2] A.Abbou, H.Mahmoudi and A.Elbacha, The effect of stator resistance variation on DTFC of induction motor and its compensation, 14th IEEE int. conference on Electronics, circuits and systems, pp.894-898, Marrakech, December 2007.

[3] N.R.N.Idris, A.M.Yatim, Direct Torque control of induction machines with constant switching frequency and reduced torque ripple, IEEE trans. on Industrial Electronics,vol.51, N°4, August 2004.

[4] J.Belhadj, I.Slama-Belkhodja, M. Pietrzak-David and B. De Fornel, Direct Torque control of induction machine with a short-time computing observer, Electromotion, Vol.10, Marrakesh, Morocco, 2003, pp.449-454.

[5] G.S.Buja and M.P.Kazmierkowski, Direct torque control of PWM inverter-fed Ac motors-Asurvey, IEEE trans. Industrial Electronics, vol.51,N°4, pp.744-757,Aug2004.

[6] L. Tang, L.Zhong, M.F. Rahman and Y.Hu, A Novel Direct Torque Control Scheme for Interior Permanent Magnet Synchronous Machine Drive system with low ripple in Torque and flux and fixed switching frequency, Conf. Rec. IEEE PESC'02 33rd Power Electronics Specialists'Conference, Cains, Australia, pp.529-534, 2002.

[7] Xianmin Ma and Zhi Na, Neural network speed identification scheme for speed sensorless DTC induction motor drive system, Power Electronics and Motion control conference IPEMC, Vol.3,pp.1242-1245,2000.

[8] Cirrincione, M.Chuan Lu and M.Pucci, Direct Torque Control of induction motors by use of the GMR neural Network, International Joint Conference on Neural Networks, Vol.3, pp.2106-2111, 2003.

[9] Microslaw Wlas, Zbigniew and Hamid A.Toliyat, Artificial-Neural Network Based Sensorless Nonlinear Control of Induction Motor,IEEE Transaction on Energy Conversion, Vol.20, No.3,2005.

[10] J. Holtz, Sensorless Control of induction motor drives, Proc. Of the IEEE, Vol. 90, No.8, 2002.

[11] Takahashi and T. Noguchi, a new quick response and high-efficiency control strategy of induction motor, IEEE Trans. On IA, Vol.22, No.5, pp. 820-827, 1987.

[12] Ji-Su Ryu, In-Sic Yoon, Kee-Sang Lee and Soon-chan Hong, Direct Torque Control of Induction Motors Using Fuzzy Variables witching Sector, IEEE International Symposium on Industrial Electronics, *ISIE*, Vol.2,pp.901-906, 2001.

[13] F.Zidani and R.N.Said, Direct torque control of induction motor with fuzzy minimization torque ripple, Journal Electrical Engineering, vol.56, pp.183-188, 2005.

[14] R.Toufouti, S.Meziane and H.Benalla, Direct Torque control for Induction Motor using Fuzzy Logic, ICGST Trans. On ACSE, Vol.6, Issue 2,pp.17-24.2006.

[15] A. Abbou and H. Mahmoudi, Sensorless Direct Torque and Flux Control of Induction Motor associated to the three levels NPC Converter Used in Electric Vehicle, International Conference on Systems and Control, CSC'2007, 16-18 May, Marrakech, 2007.

[16] N.Teske, G.M.Asher, M.Sumner and K.J. Brad ley, Suppression of Effects for the sensorless saturation saliency position control of induction motor drives under loaded condition, IEEE Tran. On Indus. Electronics, Vol. 47, No.5, pp. 1142-1149, 2000.

Optimal Colour Image Watermarking Using Neural Networks and Multiobjective Memetic Optimization

Hieu V. Dang and Witold Kinsner

***Abstract*—This paper deals with the problem of robust and perceptual logo watermarking for colour images. In particular, we investigate trade-off factors in designing efficient watermarking techniques to maximize the quality of watermarked images and the robustness of watermark. With the fixed size of a logo watermark, there is a conflict between these two objectives, thus a multiobjective optimization problem is introduced. We propose to use a hybrid between general regression neural networks (GRNNs) and multiobjective memetic algorithms (MOMA) to solve this challenging problem. Specifically, a GRNN is used for efficient watermark embedding and extraction in the wavelet domain. Optimal watermark embedding factors and the smooth parameter of the GRNN are searched by a MOMA for optimally embedding watermark bits into wavelet coefficients. The experimental results show that the proposed approach achieves robustness and imperceptibility in watermarking.**

I. INTRODUCTION

Watermarking is the technique of embedding information (watermark) into a carrier signal (video, image, audio, text) such that the watermark can be extracted or detected later for copyright protection, content authentification, identity, fingerpringing, access control, copy control, and broadcast monitoring [1]. The important requirements for the watermarking systems are robustness, transparency, capacity, and security under different attacks and varying conditions [2], [3]. These requirements can vary under different applications. Consequently, a good watermarking technique should be adaptive to the environment. A more advanced approach should involve perception, cognition, and learning [4], [5]. In general, digital watermarking can be categorized into two classes, depending on the domain of embedding the watermark [1], (*i*) spatial domain watermarking, and (*ii*) transformed domain watermarking. Digital watermarking techniques are also classified based on the watermark data embedded into the host signal. A logo watermarking technique requires a visual watermark like a logo image, while a statistical watermarking technique requires a statistical watermark like a pseudo random sequence. In statistical watermarking approaches (eg., [6], [7]), watermarks are detected by statistical method to demonstrate that the watermark in the host signal is unchanged. In logo watermarking (eg., [8], [9]), visual watermarks are extracted from the host signals for visual copyright proofs. These watermarks are not only assessed by machines but also by humans through their ability to recognize visual patterns through *human visual system* (HVS). Thus, the presentation of a visual watermark is much more persuasive than a numerical value of a statistical watermark.

Hieu V. Dang and Witold Kinsner are with the Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB, R3T 5V6, Canada. E-mails:{dangh@myumanitoba.ca|witold.kinsner@umanitoba.ca}.

Transparency and robustness are two main challenges in logo watermarking techniques since the logo consists of much information that is not easy to embed perceptually into a host signal. Moreover, the robustness in logo watermarking is so strict that it requires satisfactory recognition from human beings. With a fixed size of a logo watermark, there is a conflict between the transparency and robustness of the watermark. Increasing the transparency of watermark (or the quality of the watermarked image) decreases the robustness of the watermark and vice versa. A good logo watermarking is a robust watermarking with the acceptable quality of watermarked image. Thus, an optimal logo watermarking should be modeled as a multiobjective optimization problem.

Recently, some researchers have applied computational intelligence to design perceptual and robust watermarking systems such as *backpropagation neural networks* (BPNN) based watermarking [2], [8], [10], *support vector machine* (SVM) based watermarking [11], [12], [13], and *genetic algorithms* (GA) based watermarking [14], [15], which can detect or extract the watermark without requiring the original signal for comparison. BPNNs have been recently exploited for intelligent watermarkig methods [2], [8]. The BPNNs have been used to extract the relationships between selected pixels or selected transformed coefficients and their neighbors for embedding and extracting the watermark bits. Thus, these algorithms are robust to the amplitude scaling and a number of other attacks. However, one key disadvantage of the BPNN is that it can take a large number of iterations to converge to the desired solution [16], [18]. The watermarking problems have been recently considered as single optimization problems. Shieh and coworkers [14] introduced a watermarking technique that use a GA to find the optimum frequency bands for embedding watermark bits into *discrete cosine transform* (DCT) coefficients that can improve imperceptibility or robustness of the watermark.

In this paper, an optimal logo watermarking for colour images is formulated as a multiobjective optimization problem. To solve this problem, we propose a novel logo watermarking method based on wavelets, and the hybrid of a *general regression neural network* (GRNN) and a *multiobjective memetic algorithm* (MOMA). This new method is different from previous techniques in that it utilizes a GRNN to extract relationships between wavelet coefficiets of the Y channel of the corresponding YCrCb image for embedding and extracting the watermark. Embedding factors (watermarking strenghs) and GRNN's smooth parameter are searched optimally by a

ISBN: 978-1-61804-227-9

MOMA to maximize the quality of the watermarked image and the robustness of the watermark. The main contributions of this work are as follows:

1. A multiobjective optimization problem of logo watermarking for colour images is introduced; and
2. A novel logo watermarking method for colour images is proposed based on wavelets and GRNN. The optimality of the method is achieved by using a MOMA. This is the first MOMA based approach to optimize a logo watermarking for colour images.

The paper is organized as follows: In Sec. II, the background of GRNN is discussed. The proposed algorithms are introduced in Sec. III. Experimental results and discussions are given in Sec. IV.

II. GENERAL REGRESSION NEURAL NETWORKS

Artificial neural networks are models inspired by the working of the human brain. They are set up with some unique attributes such as universal approximation (input-output mapping), the ability to learn from and adapt to their enviroment, and the ability to invoke weak assumptions about the underlying physical phenomena responsible for the generation of the input data [16]. A neural network can provide an approximation to any function of the input vector, provided the network a sufficient number of nodes [17]. Because of those universal features, neural networks are studied extensively for applications in classification, pattern recognition, forecasting, process control, image compression, and others. Various classes of neural networks such as perceptron networks, multilayer perceptron networks, radial-basis function networks, self-organizing map networks, recurrent networks, and probabilistic networks have been proposed [16]. In this section, we will provide a brief overview of the GRNN.

The GRNN, proposed by Specht [18], is a special network in the category of probabilistic neural networks (PNN). GRNN is an one-pass learning algorithm with a highly parallel structure. Different from other probabilistic neural networks, GRNNs provide estimates of continuous variables and converges to the underlying (linear or nonlinear) regression surface. This makes GRNN a powerful tool to do predictions, approximation, and comparisons of large data sets. It also allows to have fast training and simple implementation. GRNN is sucessfully applied for image quality assessment [19], function approximation [20], and web-site analysis and categorization [21].

A diagram of the GRNN is shown in Fig. 1. In this diagram, a simple example of an one-dimensional input vector $\boldsymbol{X}[1, Q]$ is used to explain the calculation principle of the network. With the input of multidimensional vectors (i.e., matrices), it is considered as the vectors of one dimensional vector. The network has Q neurons at the input layer, Q neurons at the pattern layer, two neurons at the summation layer, and one neuron at the output layer. The input units are the distribution units. There is no calculation at this layer. It just distributes all of the measurement variable $\boldsymbol{X}$ to all of the neurons in the pattern units layer. The pattern units first calculate the cluster center of the input vector, $\boldsymbol{X}^i$. When a new vector $\boldsymbol{X}$ is entered the network, it is subtracted from the corresponding stored cluster center. The square differences d_i^2 are summed and fed into the activation function $f(x)$, and are given by

$$d_i^2 = (\boldsymbol{X} - \boldsymbol{X}^i)^T * (\boldsymbol{X} - \boldsymbol{X}^i) \tag{1}$$

$$f_i(\boldsymbol{X}) = \exp\left(-\frac{d_i^2}{2\sigma^2}\right) \tag{2}$$

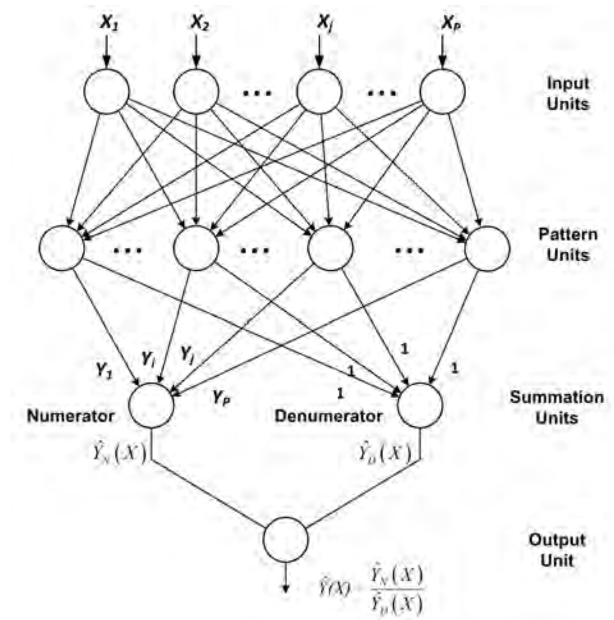


Fig. 1. GRNN block diagram.

The signal of a pattern neuron i going to the *numerator* neuron is weighted with corresponding values of the observed values (target values), Y_i, to obtain the output value of the numerator neuron, $\hat{Y}_N(\boldsymbol{X})$. The weights of the signals going to the *denumerator* neuron are one, and the output value of the denumerator neuron is $\hat{Y}_D(\boldsymbol{X})$. The output value of the GRNN is the division of $\hat{Y}_N(\boldsymbol{X})$ and $\hat{Y}_D(\boldsymbol{X})$.

$$\hat{Y}_N(\boldsymbol{X}) = \sum_{i=1}^{Q} Y_i f_i(\boldsymbol{X}) \tag{3}$$

$$\hat{Y}_D(\boldsymbol{X}) = \sum_{i=1}^{Q} f_i(\boldsymbol{X}) \tag{4}$$

The output of GRNN is given by

$$\hat{Y}(\boldsymbol{X}) = \frac{\sum_{i=1}^{Q} Y_i f_i(\boldsymbol{X})}{\sum_{i=1}^{Q} f_i(\boldsymbol{X})} \tag{5}$$

In GRNN, only the standard deviation or a smooth parameter, σ, is subject to a search. To select a good value of, σ, Specht recommends the use of the holdout method [18]. In our work, the optimal σ is searched by a multiobjective memetic algorithm for a perceptual and robust logo image watermarking.

ISBN: 978-1-61804-227-9

III. PROPOSED ALGORITHMS

A. *Watermark Embedding Algorithm*

The proposed watermark embedding scheme is depicted in the Fig. 2. In this work, we use an RGB colour image as the host image. The watermark image is a binary logo image. The RGB image is first converted to YCrCb colour image. The luminance component Y is decomposed by wavelet transform. In this paper, we only select the luminance component Y of YCbCr colour image for embedding the watermark because of the following reasons: (i) colour channels Cr and Cb have so much redundant information for HVS so that compression techniques for colour images do most compression work in these colour channels (hence, embedding watermark in CrCb will create more redundancy and make watermark susceptible to compression attacks); (ii) luminance Y is more sensitive to HVS that any tampering is easily detected (this makes watermarking in Y channel more robust than watermarking in color channels CrCb). The wavelet coefficients in each band are grouped into 3-by-3 non-overlapping blocks. Based on the random number sequence generated from the key (i, p), the algorithm selects which blocks for embedding watermark. These coefficients are used to train the GRNN. The watermark bits are embedded into selected coefficients by training the GRNN. Finally, inverse wavelet transform IDWT is applied to reconstruct the watermarked image.

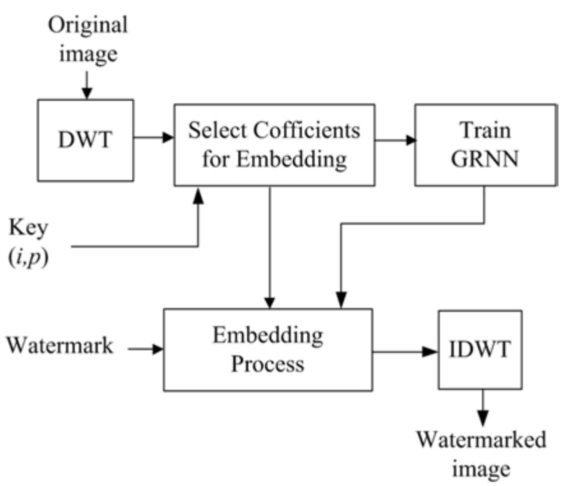


Fig. 2. Block diagram of the proposed watermark embedding scheme.

The Y component is decomposed by Symlet-2 $(sym2)$ DWT in four levels as shown in Fig. 3. The watermark bits are embedded only into the following subbands: HL^4, LH^4, HH^4, HL^3, LH^3, HH^3, HL^2, LH^2, HH^2, HL^1, LH^1. In our scheme, scaling coefficients in LL^4 and coefficients in HH^1 are not used for embedding the watermark since embedding in LL^4 will degrade the watermarked image while embedding the watermark in subband HH^1 will make the watermark more susceptible. These selected subbands are divided into non-overlapping 3-by-3 blocks and then scanned to arrange into a sequence of blocks with the subband order $HL^4LH^4HH^4HL^3LH^3HL^2LH^2HH^2HL^1LH^1$. The blocks for embedding watermarks are then selected randomly by the sequence of random non-repeated integer numbers

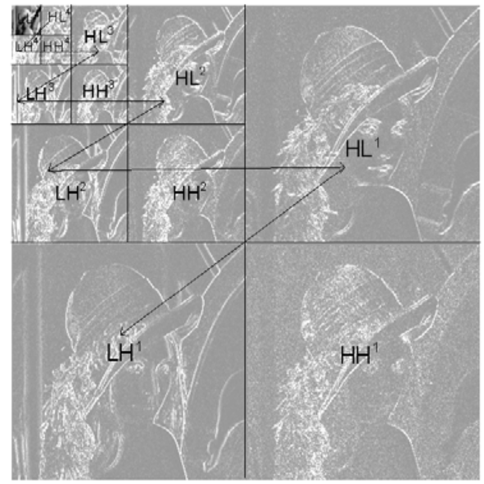


Fig. 3. Intensity-adjusted display of 4-level wavelet decomposition of Lena colour image (wavelet subbands are rescaled to a gray-intensity range for display), and the scanning order of subbands for watermarking.

generated by the Fibonacci p-code algorithm using the key (i, p). The relationship between wavelet coefficients and its neighborhoods in selected 3-by-3 blocks are extracted by a given GRNN for watermark embedding and extracting processes. The Fibonacci p-code sequence is defined by [22]

$$F_p(n) = \begin{cases} 0 & \text{if } n = 0, \\ 1 & \text{if } n = 1, \\ F(n-1) + F(n-p-1) & \text{if } n > 1, p \in Z^+ \end{cases} \tag{6}$$

Then for K sequence (k=1,2,...,K), the sequence of random integer numbers $T_k = T_1, T_2, ..., T_K$ is generated by

$$T_k = k(F_p(n) + i) \bmod F_p(n+1) \tag{7}$$

where $k = 1, 2, 3, ..., K$; $i \in [-3, 3]$ and i is an integer such that $F_p(n) + i < F_p(n+1)$. The security key or the key to generate K non repeated random integer numbers are parameters (i, p).

We now have selected blocks for embedding watermark bits. With each block B_i having the center coefficient $I(i, j)$, the iput vector X_i and target T_i are set up as in Eq. (8) to train the GRNN with 8 input neurons, 8 pattern neurons, 2 summation neurons, and 1 output neuron. Where $i = 1, 2, ..., K$; K is the number of watermark bits.

$$\begin{cases} X_i = \big[I(i-1, j-1), I(i-1, j), I(i-1, j+1), \\ \qquad I(i, j-1), I(i, j+1), I(i+1, j-1), \\ \qquad I(i+1, j), I(i+1, j+1)\big] \\ T_i \ = [I(i, j)] \end{cases} \tag{8}$$

With each pair (X_i, T_i), the GRNN produces the ouput $\hat{I}(i, j)$. The watermark bits are embedded into the selected block-center coefficients according to

$$I_w(i, j) = \hat{I}(i, j) + \eta(i)(2W(i) - 1) \tag{9}$$

where $\eta(i)$ is the watermarking factor for each embedding watermark bits to selected block-center coefficient $I(i, j)$ of

ISBN: 978-1-61804-227-9

selected block B_i. They can be altered to obtain the imperceptibility and robustness. If η is small, we get the higher quality of watermarked image, but lower level of robustness, and vice versa. This is a trade-off between the quality of the watermarked image with the robustness of watermark. $W(i)$ is the i^{th} watermark bit in the sequential watermark bits. $I_w(i,j)$, the watermarked coefficient, is obtained by replacing the central coefficient $I(i,j)$ by the combination of the output of the GRNN $\hat{I}(i,j)$ and the watermark bit $W(i)$. After embedding, an inverse DWT is performed to get the watermarked luminance Y. By combining the watermarked Y with Cr, Cb and converting to RGB, the colour watermarked image is reconstructed. This embedding algorithm is denoted as WAT-EMB procedure.

B. Watermark Extraction Algorithm

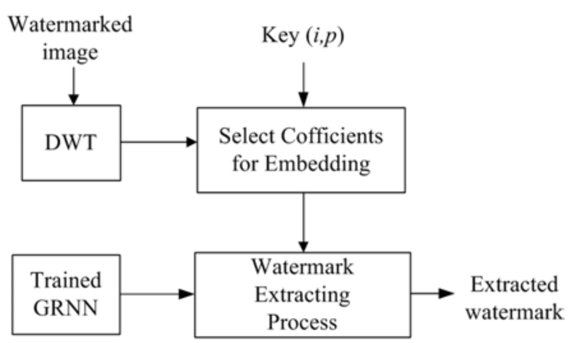


Fig. 4. Block diagram of the proposed watermark extraction scheme.

The watermark extraction scheme is illustrated in Fig. 4. The extraction process is the inverse of the embedding process. The colour watermarked image is first converted to $YCrCb$ colour domain. The luminance Y is then decomposed by 4-level Symlet-2 DWT. The wavelet coefficients are grouped into 3-by-3 blocks and arranged into the ordering sequence as described in Sec. IIA. From the key (i,p) received, the sequence of random integer numbers are generated based on the Fibonacci p-code algorithms to detect the watermarked blocks. Denote I_w is the wavelet decomposition of the component Y of the watermarked image. From the detected blocks, we setup the input vector X_i' as in Eq. (8). The trained GRNN obtained in the embedding process is used to extract the watermark bits. Each input vector X_i', the trained GRNN produce the output $\tilde{I}(i,j)$. The watermark bit extraction is performed by

$$\tilde{W}(i) = \begin{cases} 1 & \text{if } I_w(i,j) \geq \tilde{I}(i,j) \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

where i=1,2,...,K, K is the block number, and also is the number of watermark bits. $\tilde{W}$ is the extracted watermark. The extraction algorithm is denoted as WAT-EXTR procedure.

If the watermarking algorithms described in Secs. IIIA and IIIB use a fixed value η and a predifined fixed value of smooth parameter of GRNN σ (for example $\eta = 18$, $\sigma = 0.5$), we label it as WAT-GRNN algorithm.

C. Optimal Watermarking Using MOMA

In logo watemarking, with the fixed logo watemark, there always exist two conflicting objectives. These are robustness of the watermark and quality of the watermarked image (imperceptibility or transparency of watermark). In this work, we apply MOMA to search for the optimal watermarking parameters. They are the smooth parameter of the GRNN σ, and K watermarking factors $\eta(i)$, $i = (1, 2, ..., K)$.

In MOMA, the performance not only involves the evolutionary framework, but also depends on the local search. The best trade-off between a local search and the global search provided by evolution is the foremost issue in MOMA [29]. There are different MOMA frameworks introduced in the literature for domain-specific applications [30], [31]. Ishibuchi et al. [23] introduced a MOMA framework for combinatorial optimization problems. This work adopts a hybridization of the multiobjective genetic algorithm NSGA-II introduced by Deb and coworkers [24] and a local search to produce a MOMA for the Knapsack combinatorial optimization problem. In this work, a local search is employed to refine the offsprings with a weighted sum-based scheme. The selection criterion are based on Pareto ranking and crowding distance sorting used in NSGA-II. Motivated by the work of Ishibuchi et al. [23], we proposed an optimal watermarking method using MOMA. The pseudocode of the proposed algorithm is described in Algorithm 1.

Algorithm 1 WAT-MOMA

procedure WAT_MOMA(I, W, N, i, p, p_{ls})
  Generate Random Integer Numbers RN from key (i,p)
  Generate Random Population P size N
  $P \leftarrow$ OBJ-EVAL(P, W, I, RN) ▷ Evaluate Objectives
  Fast Non-Dominated Sort
  Crowding Distance Assignment
  $itrs \leftarrow 0$
  repeat
    $itrs \leftarrow itrs + 1$
    Generate Offspring Population P_{offs}
    $P_{offs} \leftarrow$ OBJ-EVAL(P_{offs}, W, I, RN)
    $P_{impr} \leftarrow$ LOCAL-SEARCH($P_{offs}, p_{ls}, I, W, RN$)
    $P_{inter} \leftarrow P \cup P_{offs} \cup P_{impr}$
    Fast Non-Dominated Sort
    Crowding Distance Assignment
    Update Population: $P \leftarrow$ Selection(P_{inter})
  until $itrs \geq MaxItrs$
  $S_{best} \leftarrow$ Sol-Select(P)
  $I_W \leftarrow$ WAT-EMB($BSOL, I, W, RN$)
  return I_W
end procedure

The inputs consist of the N number of chromosomes in population P, the colour image I, the watermark W, key (i,p), and the probability of the local search p_{ls}. From the key (i,p), the algorithm generates a sequence of random numbers RN based on the Fibonacci p_code algorithm from Eqs. (6) and (7). Each chromosome consists of $(1+K)$ genes. The first genes represents for the smooth paramenter σ of the GRNN used for embedding and extracting the watermark. The next K genes represents K embedding factors $\eta(i)$ with $i = 1, 2, ..., K$, where K is the number of watermark bits embedded into the image. The procedure OBJ-EVAL is used

ISBN: 978-1-61804-227-9

to evaluate objectives for each chromosome in the given population. In this work, we search for optimal watermarking parameters to maximize the quality of watermarked image, and the averaged robustness of watermark in the case of noise addition attack, JPEG compression attack, amplitude scaling attacks, and filtering attacks.

The procedures "Fast Non-Dominated Sort", "Crowding Distance Assignment" are parts of the NSGA-II described in details in [24], [28]. The procedure "Generate Offspring Population" is genetic operation procedure consisted of crossover and mutation operations. In this application, we use the real-coded crossover algorithm with probability p_x, and real-coded mutation with probability p_m [25], [24]. The offsprings are refined by the Tabu local search with probability of p_{ls}. In the Tabu local search, we use weighted-sum fitness with random normalized weights introduced by [27]. The Tabu local search procedure is performed only on the best individuals of a given offspring generation. Firstly, a random weight vector is generated by [27]. Based on the generated random weights, the initial solution for local search is selected from offspring population using tournament selection with replacement. The same random weights are then used for the local search to produce improved population P_{impr} from selected initial indivisual. The intermediate population P_{inter} is created by combining the current population P, the offspring population P_{offs}, and the improved population P_{impr}. The non-dominated population P is finally updated by the selection with replacement based on the Pareto ranks and crowding distances. The algorithm finishs when it meets certain terminated conditions such as predefined number of iterations.

The best solution or best chromosome (S_{best}) will be selected from the non dominated population P. Finally, we obtained the watermarked image I_W by implementing the watermark embedding algorithm presented in Sec. IIIA (WAT-EMB) with smooth parameter $\sigma = S_{best}(1)$, embedding factors $\eta(i) = S_{best}(i+1)$, $i = 1, 2, ..., K$. At the decoder side, the watermark is extracted by the watermark extraction process presented in Sec. IIIB (WAT-EXTR). The initialization and objective evaluation algorithms are discussed as follows.

1) *Initialization*: Each chromsome represents $1 + K$ real nonegative parameters to be searched. The first parameter is smooth parameter of the GRNN, which is set in the range from 0.1 to 5. The K remaining parameters represents for the K watermarking factors $\eta(i), i = 1, 2, ..., K$. The watermarking factors are searched in a wide range from 1 to 50.

2) *Objective Function Evaluation*: In literature, the objective function is also called the fitness function. The objective function uses the *peak signal to noise ratio* (PSNR) as the quality objective, and the averaged *watermark accuracy ratio* (WAR) in the cases of four different attacks as robustness objectives. The PSNR is defined by

$$\text{PSNR} = 10 \log_{10} \left(\frac{I_{peak}^2}{\text{MSE}} \right) \tag{11}$$

where I_{peak} is the maximum intensity value of the three color channels R, G, B, and the *mean squared error* (MSE) computed for all three color channels R, G, and B is given by

$$\text{MSE} = \frac{1}{KMN} \sum_{k=1}^{3} \sum_{i=1}^{M} \sum_{j=1}^{N} (I(i,j,k) - I_W(i,j,k))^2 \tag{12}$$

The watermark accuracy ratio is defined by

$$\text{WAR} = \frac{\sum_{i=1}^{M_w} \sum_{j=1}^{N_w} W(i,j) \bar{\oplus} \tilde{W}(i,j)}{M_w * N_w} \tag{13}$$

where W and $\tilde{W}$ are the original and extracted watermarks, and (M_w, N_w) is the size of the watermarks. The logic operator $\bar{\oplus}$ do comparison between W and $\tilde{W}$. $W(i,j) \bar{\oplus} \tilde{W}(i,j) = 1$ if $W(i,j)$ and $\tilde{W}(i,j)$ have the exactly same value of 0 or 1. If WAR $\geq 70\%$, the extracted watermark can be considered as the original watermark. It is close to be perfect if WAR $\geq 85\%$.

Let $K = M_w * N_w$ be the number of watermark bits embedded into the image. We denote $\bar{\alpha} = [\alpha_1, \alpha_2, ..., \alpha_{K+1}]$ as the watermarking parameters to be searched, where $\alpha_1 = \sigma$ (the smooth parameter of the GRNN), $\alpha_{2:K+1} = \eta(1:K)$ (the embedding factors). The objectives function is then set up as follows

$$\bar{f}(\bar{\alpha}) = [f_1(\bar{\alpha}), f_2(\bar{\alpha})] \tag{14}$$

where

$$f_1(\bar{\alpha}) = \text{PSNR}(\bar{\alpha}) = \text{PSNR}(\alpha_1, \alpha_2, ..., \alpha_{K+1})$$

and

$$f_2(\bar{\alpha}) = \frac{W_G(\bar{\alpha}) + W_J(\bar{\alpha}) + W_A(\bar{\alpha}) + W_M(\bar{\alpha})}{4}$$

where W_G is the WAR in the case that the watermarked image is tampered by the Gaussian noise addition attack; W_J is the WAR under JPEG compression attack; W_A is the WAR under the amplitude scaling attack; and W_M is the WAR under the median filtering attack. Our optimal watermarking problem is to search optimal parameters $\bar{\alpha}$ that can be formed by

$$\max_{\bar{\alpha}} \bar{f}(\bar{\alpha}) = \max_{\bar{\alpha}} [f_1(\bar{\alpha}), f_2(\bar{\alpha})] \tag{15}$$

3) *Local Search*: In this work we employ the principle of Tabu local search [26] with random normalized weights generated from [27]. The best initial solution for the local search is selected by doing a tournament selection between chromosomes in the population P_{offs}. The procedure finally returns the N_{LS} better solutions, P_{impr}.

4) *Crossover, Mutation and Selection with Replacement Operations*: Genetic operators including crossover and mutation are used to generate offspring population in each evolutionary loop. In this work, the real-coded crossover and mutation introduced in [25], [24] are adopted with crossover probability $p_x = 0.8$ and mutation probability $p_m = 0.05$. The non-dominated chromosomes are selected in each evolutionary loop by using the selection with replacement based on the Pareto ranks and crowding distances as described in [24], [28].

IV. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, experimental results are demonstrated and discussed to show the watermark robustness and transparency of the proposed algorithm. In the embedding process, the memetic algorithm is used to search for optimal watermarking factors and the optimal smooth parameter of the GRNN. In the watermark extraction process, the original image is not required, but the secret key (i, p), the smooth and weight parameters of the trained GRNN from the embedding process are needed. The watermark extraction process is the same as the watermark extraction algorithm described in the Sec. IIIB (WAT-EXTR). The experimental results obtained from the proposed algorithm using multiobjective memetic algorithm (WAT-MOMA) are compared with results of the WAT-GRNN algorithm, Kutter's method [32], and Yu's method [8]. WAT-GRNN is the watermarking algorithm used WAT-EMB in Sec. IIIA and WAT-EXTR in Sec. IIIB with the fixed embedding factor (embedding strength) $\eta = 18$, and the smooth parameter of the GRNN $\sigma = 0.5$. In the Yu's and Kutter's methods, we setup the watermark strength $\alpha = 0.2$ to have a good robustness to be compared to the proposed algorithm WAT-MOMA.

To evaluate the performance of our watermarking algorithms, the "Winipeg Jet" logo is embedded into various colour images. The binary watermark of size 64-by-64 is embedded into highly-textual colour images "Lena", "Baboon", "Airplane-F16", and "House" each with size of (512-by-512)-by-3.

A. Results of Multiobjective Memetic Optimization Algorithm

In the WAT-MOMA algorithm, which uses the multiobjective memetic optimization to search for optimal watermarking factors and the smooth parameter of GRNN, the number of initial chromosomes N setup to 100, the local search is applied to refine the offspring population with the probability of 0.5 and the number of iterations is 50. These local search parameters are selected based on the analystical results shown in [23] for memetic algorithm using weighted sum-based local search. The numerical results in Fig. 5. shows that the algorithm based on memetic optimization is more effective than the algorithm based on multiobjective genetic algorithm NSGA-II [24]. Since there is conflict between the quality of watermarked image and the robustness of watermark in watermarking, the optimally selected chromosome (solution) is a balance between the PSNR objective and the averaged WARs objectives. The solution includes the smooth parameter of GRNN and 64x64=4096 embedding factors. Example of the optimal embedding factors for Lena colour image after 100 iterations are illustrated in Fig. 6 corresponding the smooth parameter of GRNN $\sigma = 2.47657$.

B. Quality Evaluation

To measure the transparency or the similarity of the watermarked image to the original image, watermarking systems mostly employ the PSNR. In Fig. 7, the differences between the original images and the watermarked images are difficult to observe by human eyes. The PSNRs obtained by WAT-MOMA for all these four colour test images are compared with PSNRs obtained by WAT-GRNN, Yu's method, and Kutter's method. The comparison results are described in Table I.

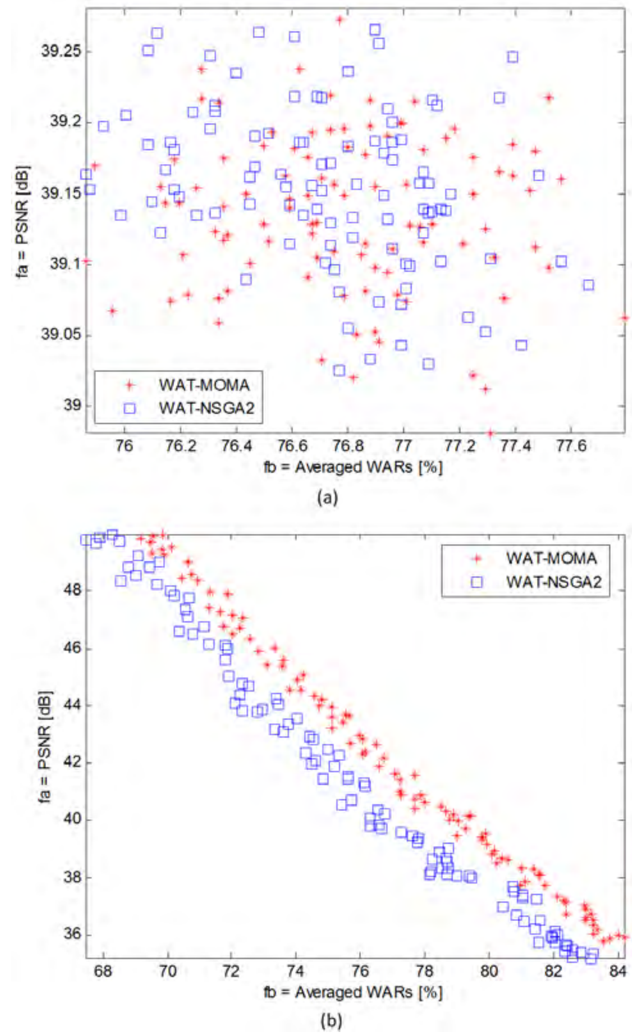


Fig. 5. Numerical results of the watermarking based on memetic and NSGA2 strategies for Lena color image: (a) WAT-MOMA's initial population versus WAT-NSGA2's, (b) WAT-MOMA's population versus WAT-NSGA2's population after 100 iterations.

TABLE I
PSNR COMPARISON OF WATERMARKED IMAGES

Images	PSNR [dB]			
	Kutter's	Yu's	WAT-GRNN	WAT-MOMA
Lena	41.8433	41.6670	42.4590	42.8180
Baboon	41.3612	41.2206	42.5781	42.4320
Airplane	38.6961	38.5295	42.3353	42.8027
House	39.4374	39.2806	42.3143	42.8596

C. Robustness Evaluation

The robustness of the watermark is evaluated by the similarity between the extracted watermark and the original wa-

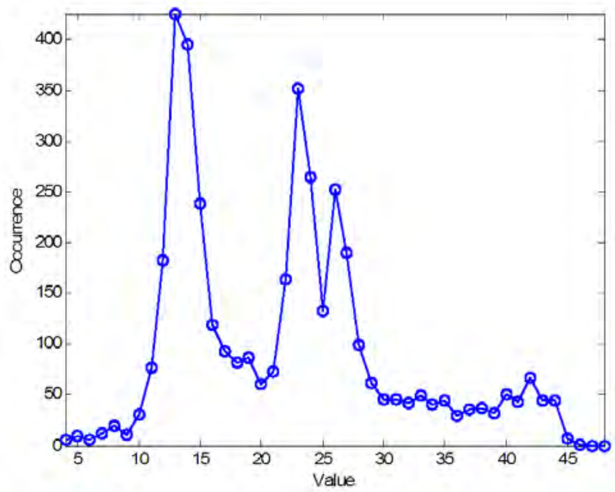


Fig. 6. Watermarking factors for Lena colour image obtained after WAT-MOMA run 100 iterations

termark throuth WAR computed by Eq. (13). The watermarks extracted from the watermarked images in Fig. 7 are shown in Fig. 8. The calculated WARs indicate that our method perfectly extracts watermarks from watermarked images in the case of without any attacks.

We test the proposed algorithm with five different classes of attacks such as (*i*) compression attacks (JPEG compression), (*ii*) noise addition attacks (AWGN, salt & pepper, and fractional noises), (*iii*) filtering attacks (median filtering), (*iv*) amplitude scaling attacks, (*v*) and geometric manipulation attacks (image cropping, and rotation). Due to space limitation, we present only some results in this section.

1. *Robustness Against JPEG Compression*: JPEG is common image compression standard for multimedia application. Hence, watermarking systems should be robust to this attack. Figure 9 shows an example of JPEG compression attack with the quality factor of 40 to the watermarked images of Lena and Baboom, and the proportional extracted watermarks. The robustness comparison with WAT-GRNN, Yu's and Kutter's methods for the watermarked image of Lena in Fig. 7 is displayed in Fig. 10.

2. *Robustness Against Amplitude Scaling*: The colour values of the watermarked image are divided by a *scaling factor* (SF). The attack is called negative amplitude scaling attack if SF is greater than one, and vice versa is the positive amplitude scaling attack. An example of the positive amplitude scaling attack with SF$=0.3$ for watermarked images of Lena and Baboom in Fig. 7 are depicted in Fig. 11. The robustness of watermark compared with results from WAT-GNRR, Yu's and Kutter's methods is illustrated in Fig. 12.

It can be seen that the WAT-MOMA algorithm is very robust to amplitude scaling attacks. Even if with the positive attack of SF=0.3 that decreases the SNR of the attacked watermarked image to -7.36 dB, we are still able to recover the watermark excellently.

3. *Robustness Against Additive White Gaussian Noise*: Since the natural features of electronic devices and communications channels, AWGN is perhaps the most common noise in

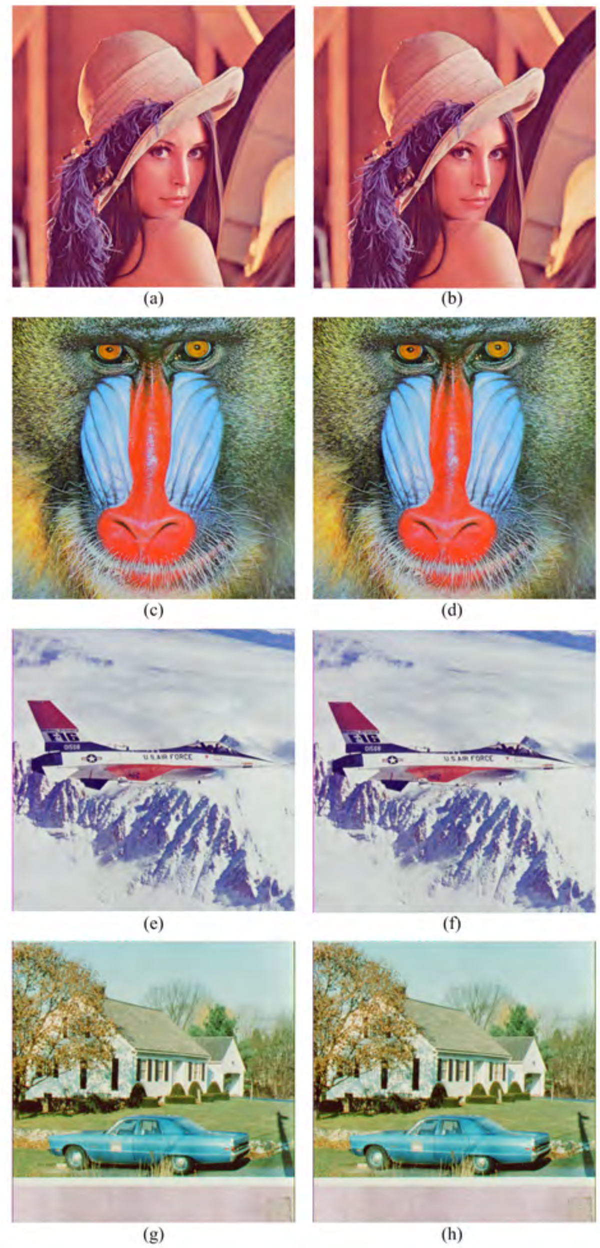


Fig. 7. The original test images and watermarked test images: (a) original Lena image, (b) watermarked lena image with the obtained PSNR=42.82 dB, (c) original Baboon image, (d) watermarked Baboon image with the obtained PSNR=42.43 dB, (e) original Airplane F16 image, (f) watermarked Airplane F16 image with the obtained PSNR=42.80 dB, (g) original House image, (h) watermarked House image withe the obtained PSNR=42.86 dB.

communications systems. Thus, a good watermarking scheme should be robust to AWGN. The robustness fo our scheme against AWGN is shown in Fig. 13 and Fig. 14.

The AWGN is added to the watermarked images with different standard deviation σ_n (corresponding SNRs). The Gaussian noise is added to the colour image of watermarked image, I_W, by

$$I_W^N = I_W + \sigma_n N \tag{16}$$

where N is the normally distributed random noise, and I_W^N is the watermarked image corrupted by the Gaussian noise.

ISBN: 978-1-61804-227-9

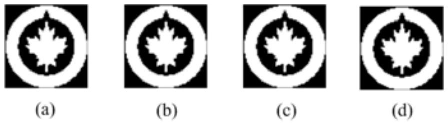


Fig. 8. Watermarks extracted from watermarked images in Fig. 7: (a) extracted from Fig. 7(b) with WAR=100 %, (b) extracted from Fig. 7(d) with WAR=100 %, (c) extracted from Fig. 7(f) with WAR=100 %, (d) extracted from Fig. 7(h) with WAR=100 %.

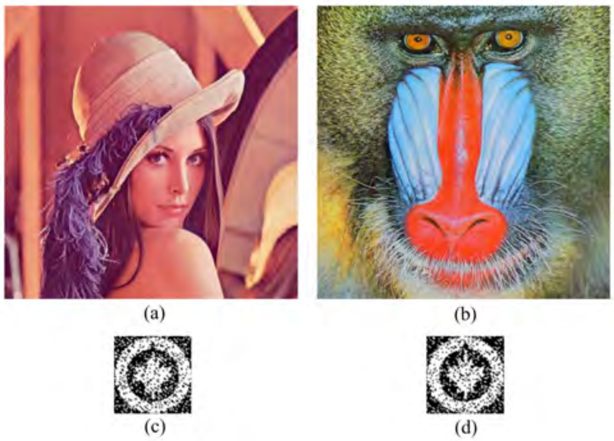


Fig. 9. An example of JPEG compression attack and watermark extraction with JPEG quality factor of 40: (a) compression of watermarked image of Lena at Fig. 7(b) with SNR=26.14 dB, (b) compression of watermarked image of Baboom at Fig. 7(d) with SNR=18.98 dB, (c) the extracted watermark from (a) with WAR=82.47 %, (d) the extracted watermark from (b) with WAR=83.42 %.

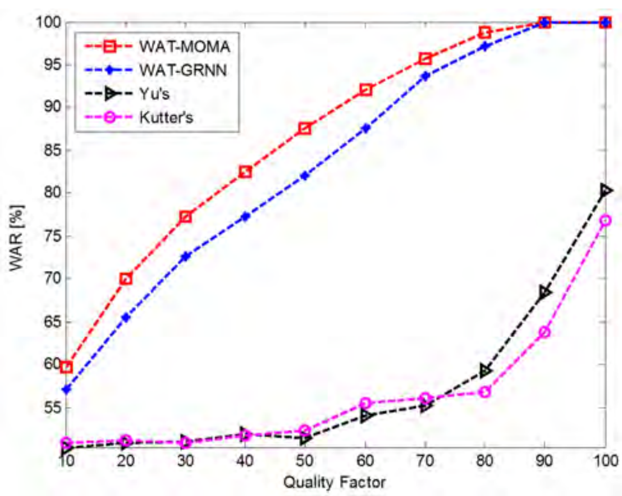


Fig. 10. The experimental results under the JPEG compression attack for watermarked image of Lena.

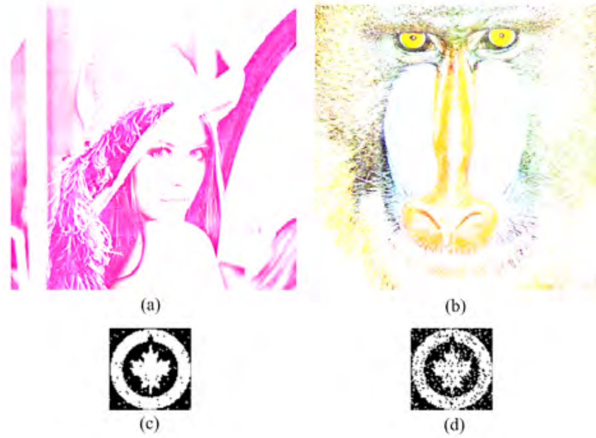


Fig. 11. An example of amplitude scaling attack and watermark extraction with SF=0.3: (a) scaling the watermarked image of Lena at Fig. 7(b) with SNR=-7.36 dB, (b) scaling the watermarked image of Baboom at Fig. 7(d) with SNR=-7.36 dB, (c) the extracted watermark from (a) with WAR=98.09 %, (d) the extracted watermark from (b) with WAR=90.09 %.

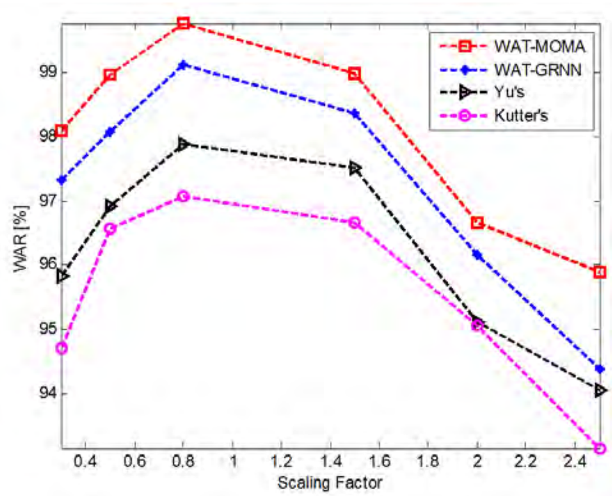


Fig. 12. The experimental results under the amplitude scaling attack for watermarked image of Lena.

The proposed method works really well, even with a variance of AWGN=40^2 (with the equivalent SNR around 10 dB). This level is a challenge to every watermarking and denoising techniques [33], [34].

4. *Robustness Against Median Filtering*: Median filtering is always a serious challenge to watermarks. This is because a median filter does average pixel values in the window size that eliminates high dynamic values in the image in the spatial domain. Hence, median filtering can affect the watermark severely. An example of doing median filtering for watemarked images of Lena and Baboon with the filter window size of 5 is displayed in the Fig. 15. The robustness comparison of the proposed algorithm with other methods for the watermarked image of Lena is depicted in Fig. 16.

V. CONCLUSIONS

In this paper, a logo watermarking for colour images is formulated as a multiobjective optimization problem of finding the watermarking parameters to maximize the quality of watermarked image and the robustness of the watermark under different attacks. A novel intelligent and robust logo watermarking method based on the general regression neural networks and multiobjective memetic algorithms is proposed to solve this challenging problem. Specifically, the embedding factors and the smooth parameter of the GRNN are searched optimally by the multiobjective memetic optimization

Fig. 13. An example of AWGN noise attack and watermark extraction with variance of AWGN= 40^2: (a) attacked watermarked image of Lena at Fig. 7(b) with SNR= 10.9 dB, (b) attacked watermarked image of Baboom at Fig. 7(d) with SNR=10.74 dB, (c) the extracted watermark from (a) with WAR=75.34 %, (d) the extracted watermark from (b) with WAR=71.73 %.

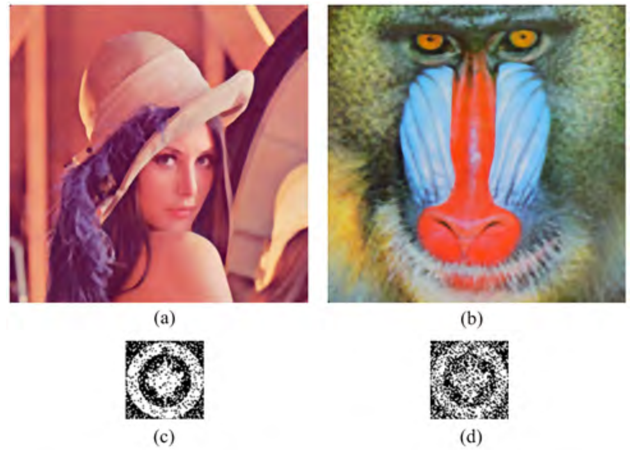


Fig. 15. Median filtering attack to the watermarked images with filter window = 5: (a) filtered watermarked Lena image with SNR=26.87dB; (b) filtered watermarked Baboon image with SNR=17.43 dB; (c) watermark extracted from (a) with WAR=84.45 %; (d) watermark extracted from (b) with WAR=70.73 %.

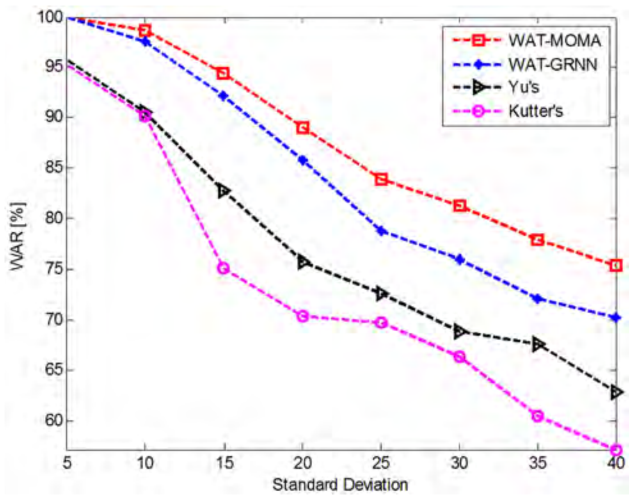


Fig. 14. The experimental results under the AGWN noise attack for watermarked image of Lena.

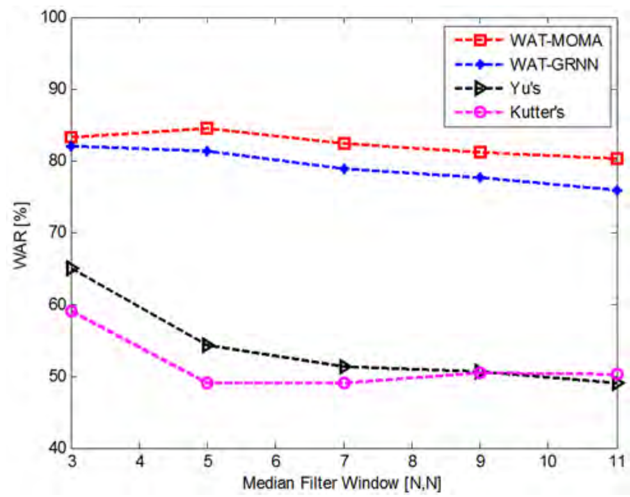


Fig. 16. Experimental results under median filtering attacks for the watermarked images of Lena.

to maximize the PSNR and the averaged WARs objectives. The proposed algorithm obtains better results in transparency and robustnesses against classes of additive noise, and signal processing attacks than previous approaches.

We discuss the application of neural networks for watermarking systems. We evaluated neural networks, and selected GRNN for its good fit to our problem. The GRNN is much superior over the BPNN when solving this problem as it has very fast time convergence and high prediction accuracy.

However, the proposed algorithm has its own disadvantages and needs further improvements. For example, since it needs a sufficient time for the evolutionary and local refining searches to find the best local and global solutions, it is not fast enought for the real-time applications at this stage.

References

[1] Min Wu and Bede Liu, "Data hiding in image and video: Part I - Fundamental issues and solutions," *IEEE Trans. Image Processing*, vol. 12, no. 06, pp. 685-695, June 2003.

[2] Jeng-Shyang Pan, Hsiang-Cheh Huang, and Lakhmi C.Jain, *Intelligent Watermarking Techniques*, New Jersey, MA: World Scienstific, 2004, 852 pp.

[3] Benoit Macq, Jana Dittmann, and Edward J. Delp, "Benchmarking of image watermarking algorithms for digital rights management," *Proc. IEEE*, vol. 92, no. 6, June 2004.

[4] Witold Kinsner, *Towards cognitive security systems*, in *Proc. of the 11th IEEE Intern. Conf. on Cognitive Informatics and Cognitive Computing*, ICCI*CC 2012, (Kyoto, Japan; August 22-24, 2012), 2012, (Keynote Speech).

[5] Yingxu Wang, James A. Anderson, George Baciu, Gerhard Budin, D. Frank Hsu, Mitsuru Ishizuka, Witold Kinsner, Fumio Mizoguchi, Kenji Sugawara, Shusaku Tsumoto, Du Zhang, "Perspectives on eBrain and cognitive computing," *International Journal of Cognitive Informatics and Natural Intelligence* (IJCINI), vol. 6, no. 4, pp. 1-21, Oct-Dec 2012.

[6] Mauro Barni, Franco Bartolini, and Alessandro Piva, "Improved wavelet-based watermarking through pixel-wise masking," *IEEE Trans. Image Processing*, vol. 10, no. 5, pp. 783-791, May 2001.

[7] Prayoth Kumsawat, Kitti Attakitmongcol, and Arthit Srikaew, "A new approach for optimization in image watermarking by using genetic algorithms," *IEEE Trans. Sig. Processing*, vol.53, no.12, Dec. 2005.

[8] Pao-Ta Yu, Hung-Hsu Tsai, and Jyh-Shyan Lin, "Digital watermarking based on neural networks for color images," *Signal Processing*, vol. 81, no. 3, pp. 663-671, March 2001.

ISBN: 978-1-61804-227-9

[9] Chun-Hsien Chou and Kuo-cheng Liu, ”A perceptually tuned watermarking scheme for color images,” *IEEE Trans. Image Processing*, vol. 19, no. 11, pp. 2966-2982, Nov. 2010.

[10] Hieu V. Dang and Witold Kinsner, ”An intelligent digital color image watermarking approach based on wavelet transform and general regression neural network,” in *Proc. of the 11th IEEE Intern. Conf. on Cognitive Informatics and Cognitive Computing*, ICCI*CC 2012, (Kyoto, Japan; August 22-24, 2012), pp. 115-123, 2012.

[11] Xiang-Yang Wang, Hong-Ying Yang, and Chang-Ying Cui, ”An SVM-based robust digital image watermarking against desynchronization attacks,” *Signal Processing*, vol. 88, no. 9, pp. 2193-2205, Sep. 2008.

[12] Hung-Hsu Tsai and Duen-Wu Sun, ”Color image watermark extraction based on support vector machines,” *Information Sciences*, vol. 177, no. 2, pp.550-569, Jan. 2007.

[13] Rui-min Shen, Yong-gang Fu, and Hong-tao Lu, ”A novel image watermarking scheme based on support vector regression,” *The Journal of Systems and Software*, vol. 78, no. 1, pp. 1-8, Oct. 2005.

[14] Chin-Shiuh Shieh, Hsiang-Cheh Huang, Feng-Hsing Wang, and Jeng-Shyang Pan, ”Genetic watermarking based on transform-domain techniques,” *Pattern Recognition*, vol. 37, no. 3, pp.555-565, March 2004.

[15] K. Ramanjaneyulu and K. Rajarajeswari, ”Wavelet-based oblivious image watermarking scheme using genetic algorithm,” *IET Image Process.*, vol. 6, no. 4, pp. 364-373, June 2012.

[16] Simon Haykin, *Neural networks: A comprehensive foundation*, Second Edi. Pearson, 1999, 823 pp.

[17] L. Marquez and T. Hill, ”Function approximation using backpropagation and general regression neural networks,” in *Proceedings of the 16th IEEE Int. Conf. System Sciences*, Hawaii, pp. 607-615, 1993.

[18] Donald F. Specht, ”A general regression neural network,” *IEEE Trans. Neural Networks*, vol. 2, no. 6, pp. 568-576, Nov. 1991.

[19] Chaofeng Li, Alan C. Bovik, and Xiaojun Wu, ”Blind image quality assessment using a general regression neural network,” *IEEE. Trans. Neural Netw.*, vol. 22, no. 5, pp. 793-799, May 2011.

[20] John Y. Goulermas, Panos Liatsis, Xiao-Jun Zeng, and Phil Cook, ”Density-driven generalized regression neural networks (DD-GRNN) for function approximation.” *IEEE Trans. Neural Netw.*, vol. 18, no. 6, pp. 1683-1696, Nov. 2007.

[21] Ioannis Anagnostopoulos, Christos Anagnostopoulos, George Kouzas, and Dimitrios D. Vergados, ”A generalised regression algorithm for web page categorisation,” *Neural Comput. Appl.*, vol. 13, no. 3, pp. 229-263, Sep. 2004.

[22] Yicong Zhou, Sos Agaian, Valencia M. Joyner, and Karen Panetta, ”Two Fibonacci p-code based image scrambling algorithms,” *SPIE Proceedings: Image processing - algorithms and systems VI*, vol. 6812, Paper #6812-15, San Jose, CA, January 2008.

[23] Hisao Ishibuchi, Yasuhiro Hitotsuyanagi, Noritaka Tsukamoto, and Yusuke Nojima, ”Implementation of multiobjective memetic algorithms for combinatorial optimization problems: A Knapsack problem case study,” in *Multi-Objective Memetic Algorithms*, Chi-Keong Goh, Yew-Soon Ong, and Kay Chen Tan, (eds.) Berlin Springer, pp. 27-49, 2009.

[24] Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and T. Meyaruvan, ”A fast and elitist multiobjective genetic algorithm: NSGA-II,” *IEEE Trans. Evol. Comp.*, vol. 6, no. 2, pp. 182-197, April 2002.

[25] K. Deb and R. B. Agrawal, ”Simulated binary crossover for continuous search space,” Complex Syst., vol. 9, pp. 115-148, Nov. 1995.

[26] Fred Glover, ”Tabu search: a tutorial,” Interfaces, vol. 20, no. 4, pp. 74-94, Aug. 1990.

[27] Andrzej Jaszkiewicz, ”Gnetic local search for multi-objective combinatorial optimization,” *European Journal of Operational Research*, vol. 137, no. 1 pp.50-71, Feb. 2002.

[28] K.C.Tan, E.F.Khor, and T.H.Lee, *Multiobjective evolutionary algorithms and applications*, London UK: Springer, 2010, 295 pp.

[29] Natalio Krasnogor and Jim Smith, ”A tutorial for competent memetic algorithms: model, taxonomy, and design issues,” *IEEE Trans. Evol. Comput.*, vol. 9, no. 5, pp. 474-488, Oct. 2005.

[30] Daniel Molina, Manuel Lozano, Carlos Garcia-Martinez, and Francisco Herrera, ”Memetic algorithms for continuous optimisation based on local search chains,” *Evolutionary Computation*, vol. 18, no. 1, pp. 27-63, 2010.

[31] Andriana Lara, Gustavo Sanchez, Carlos A. Coello, and Oliver Schutze, ”HCS: A new local serach strategy for memetic multiobjective evolutionary algorithms,” *IEEE Trans. Evol. Comput.*, vol. 14, no. 1, pp. 112-132, Feb. 2010.

[32] M. Kutter, F. Jordan, and F. Bossen, ”Digital watermarking of color images using amplitude modulation,” *Joural of Electronic Imaging*, vol. 7, no. 2, pp. 326-332, 1998.

[33] David L. Donoho, ”De-noising by soft-thresholding,” *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613-627, May 1995.

[34] Witold Kinsner, ”Compression and its metrics for multimedia,” in *Proc. of the 11th IEEE Intern. Conf. on Cognitive Informatics*, ICCI 02, (Calgary, Canada; August 19-20, 2002), 2002.

Automatic Generation of Fuzzy Classification Rules from Data

Mohammed Al-Shammaa[1] and Maysam F. Abbod[2]

***Abstract*—**In this paper, we propose a method for automatic generation of fuzzy rules for data classification. The proposed method is based on subtractive clustering optimized using genetic algorithm. It searches for the FIS structure and number of rules that have the highest fitness value. Multiple performance measures are incorporated into the fitness function to address the problem of imbalanced data. Fitness function includes both training and validation to avoid data over-fitting. Classification performance of the proposed method is evaluated using different data sets and results are compared to those of a number of models generated by fuzzy c-means clustering with various cluster numbers. Results show that the proposed method has better accuracy and a well compromised sensitivity and specificity.

***Keywords*—**Fuzzy rules generation, Genetic algorithm, Neuro-fuzzy system, Subtractive clustering.

I. Introduction

Fuzzy inference system (FIS) is a computational technique that relies on fuzzy logic [1] for performing input-output mapping. FIS is featured by its interpretability, ability to deal with vagueness and incompleteness of knowledge and the ability of describing non-linear relationships between input and output. However, it lacks learning capability and systematic design [2].

One of the most important issues of designing FIS is determining the structure of its rule base that best suits the problem. In most classification problems, it is difficult for human experts to provide sufficient information required to generate fuzzy rules. Therefore, computational methods are used to generate fuzzy rules from data automatically.

Several methods for extracting fuzzy rules from data have been proposed. Data clustering is one of the most widely used approaches for rule base generation. Data clustering methods aim to group data into clusters based on a similarity measure. Fuzzy sets of an input variable can be obtained from data clusters by projecting each cluster to the dimension that corresponds to the input. For example, subtractive clustering method presented in [3] is used to extract fuzzy rules from data. This method selects the data point with many neighboring data points as a cluster center and associates the neighboring data points to this cluster. Another clustering method used to generate fuzzy rules is fuzzy clustering [4]. In this method, data elements can be associated to more than one cluster. Each data point is assigned a set of membership levels which indicate the strength of its association to one or more clusters.

The use of optimization algorithms for fuzzy rules generation is suggested by many studies. For example, genetic algorithm is used to simultaneously estimate rule base structure and parameters of the fuzzy model from data [5]. Also, particle swarm optimization is used to obtain the antecedents, consequences, and connectives of the fuzzy rules base [6]. The use of fuzzy genetic algorithm for automatic fuzzy rule generation is investigated in [7].

In addition to data clustering and optimization algorithms, several methods are suggested by the literature. A method presented in [8] aims to generate fuzzy rules automatically by optimizing fuzzy neural network. The optimization is performed by a hybrid algorithm based on tabu search algorithm and least squares algorithm. Algorithms based on a tolerance rough sets model are used in [9] to obtain fuzzy rules. Resulted rules are then fuzzified using a genetic algorithm.

A method that uses unsupervised learning and reinforcement learning for structure identification and parameter estimations of fuzzy inference systems is suggested in [2]. In this method, unsupervised learning clustering methods to cluster data and generate fuzzy inference systems while adjustment and deletion of fuzzy rules are achieved using reinforcement learning. In [10], a data mining approach based on regularization theory is used to refine and generate fuzzy classification rules.

In this paper, we present a method for automatic generation of fuzzy classification rules from data. The proposed method is based on subtractive clustering optimized using genetic algorithm.

This paper is arranged as follows. A brief description of fuzzy inference system is given in section II. Section III introduces genetic algorithms. The proposed method is presented in section IV. Section V shows experimental results obtained by testing the proposed method. The conclusion is given in section VI.

1 School of Engineering and Design, Brunel University, London, UB8 3PH, UK; E-Mail: Mohammed.Al-Shammaa@brunel.ac.uk

PhD student, sponsored by the Higher Committee for Education development, Iraq.

2 School of Engineering and Design, Brunel University, London, UB8 3PH, UK; E-Mail: maysam.abbod@brunel.ac.uk

II. Fuzzy Inference System

Fuzzy inference systems (FIS) are widely used in many applications, from system modelling, simulation and control to classification and decision support. FIS is a computational technique that relies on fuzzy logic [1] for performing input-output mapping. It inherits, from fuzzy logic, the ability to deal with vagueness and incompleteness of knowledge and the ability of describing non-linear relationships between input and output.

FIS consists of three components: a set of fuzzy if-then rules, a database that defines the fuzzy membership functions used in these rules and a reasoning mechanism that performs the inference process. Inputs are transformed to fuzzy sets using the membership functions defined in the FIS's database and the output is computed from the fuzzy rules using the reasoning mechanism. The consequent part of FIS's rules can be a fuzzy value (Mamdani model [11]), a crisp value (type II), or a function of linear combination of input variables (TSK model [12]). In the case of Mamdani model, the process of defuzzification is required to produce the final crisp value of the output.

Information required to design FIS and create its rules can be obtained from the knowledge of human experts. However, with increasing the complexity of the system, the number of inputs and the number of fuzzy variables for each input, it becomes more difficult for human experts to generate these rules. Instead, some computational methods (e.g. data clustering) can be used to generate fuzzy rules from data automatically.

Data Clustering

Data clustering methods aim to group data into clusters of data points that feature a certain level of similarity more than those in other clusters. Fuzzy sets of an input variable can be obtained from data clusters by projecting each cluster to the dimension that corresponds to the input. Thus, each cluster represents a fuzzy rule whose antecedent part is the fuzzy sets resulted from this projection. Many clustering methods require the number of clusters to be specified prior to the clustering process. Different number of clusters results in different clusters properties, hence different fuzzy rules. In addition to the number of clusters, each clustering method has its own parameters that affect the clustering process. Finding the right number of clusters and other clustering parameters for a given application is difficult in many cases. On the other hand, there are some clustering methods that do not require the number of clusters to be specified, instead, they rely on a number of parameters to perform clustering. One of these methods is subtractive clustering [3]. Initially, this method considers that each point is a potential cluster center. Then, for each data point, it computes the potential, defined as the density of the neighboring data points which is a function of the Euclidean distances to all other data points:

$$P_i = \sum_{j=1}^{n} e^{-(4/r_a^2)\left\|x_i - x_j\right\|^2} \tag{1}$$

where P_i is the potential of data point i and r_a is the radius in which other points have significant effect on the potential of data point i.

The point with greatest potential is chosen as a first cluster center, and the potentials of all other points are revised by subtracting an amount that is a function of their distances to the first center:

$$P_i \Leftarrow P_i - P_1^* e^{-(4/r_b^2)\left\|x_i - x_1^*\right\|} \tag{2}$$

where r_b is the radius in which points will have significant reduction in potential, i.e. are less likely to be selected as the next cluster center. This radius can be expressed as multiplication of r_a by a factor S called squash factor:

$$r_a = S\, r_b \tag{3}$$

After potential revision, the data point with the next higher potential is selected as a second cluster center candidate. Whether this point is accepted as a new cluster center or not depends on its potential according to the following conditions:

If $P_k^* > \bar{\varepsilon} P_1^*$, x_k^* is accepted as a new cluster center and the above procedure is repeated.

If $P_k^* < \underline{\varepsilon} P_1^*$, x_k^* is rejected as a new cluster center and the clustering is ended.

If $\bar{\varepsilon} P_1^* > P_k^* > \underline{\varepsilon} P_1^*$, then if $\dfrac{d_{\min}}{r_a} + \dfrac{P_k^*}{P_1^*} \geq 1$, x_k^* is accepted as a new cluster center and the above procedure is repeated. Otherwise x_k^* is rejected, its potential set to 0 and the procedure is repeated with the data point with next higher potential. Here, $d_{\min}$ is the shortest distance between x_k^* and all previously found cluster centers and $\bar{\varepsilon}$ and $\underline{\varepsilon}$ are accept and reject ratios respectively. So, the properties of the final clusters resulted from this method depends on the following parameters: squash factor (S), accept ratio ($\bar{\varepsilon}$), reject ratio ($\underline{\varepsilon}$), and radius (r_a) (or a number of different radii equal to the dimension of the clustered data).

III. Genetic Algorithm

Genetic algorithm (GA) is a stochastic optimization algorithm that imitates natural evolution. Like in natural evolution, genetic algorithms generate populations of different possible solutions successively. Each generation consists of a number of individual solutions, called chromosomes, which compete with each other. The fittest chromosomes with respect to some criteria get higher probabilities of being selected to reproduce the next generation. A chromosome comprises a set of coded properties, called genes, which describe the solution. Genes are either binary coded or real coded [13] [14]. The next generation of chromosomes is produced from selected chromosomes, called parents, by a reproduction method. New parents are selected for each new child to be generated. There are various procedures to select the parents like: roulette

ISBN: 978-1-61804-227-9

wheel, linear ranking and geometric ranking. The reproduction of a new generation from parents is usually done using two main operators: crossover and mutation. In crossover, chromosomes of one child or more are produced from genes of two or more parent chromosomes using a crossover technique. On the other hand, mutation produces one new chromosome by altering one or more genes in a single chromosome [15].

IV. Proposed method

In this paper, we present a method that aims to generate fuzzy rules from data automatically to build a FIS used in data classification. The proposed method relies on subtractive clustering to create data clusters from which fuzzy rules can be obtained. The clustering process is optimized using genetic algorithm to produce a FIS that classify data as accurate as possible. Input membership functions of the resulted FIS are then fine-tuned using GA. The final FIS can be used as an initial FIS for neuro-fuzzy system. The block diagram of the proposed method is shown in Fig. 1. It comprises of two main parts: subtractive clustering and genetic algorithm.

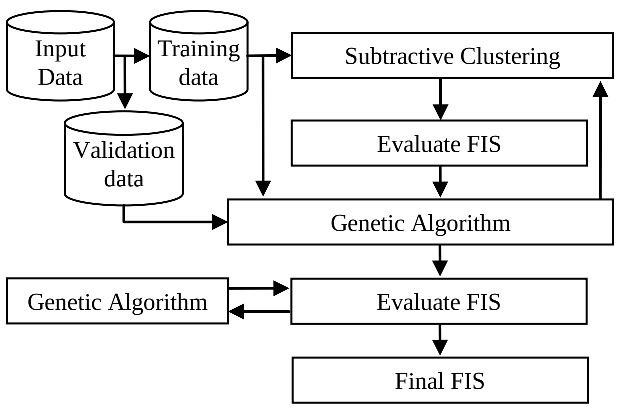


Fig.1. Block diagram of proposed method.

Modeling data are divided into two sets: training data (70%) and validation data (30%).

The parameters of subtractive clustering algorithm (squash factor, accept ratio, reject ratio, and a number of radii equal to the dimension of the clustered data) are adjusted by the GA. Each chromosome contains a set of possible values for these parameters. Table I depicts the structure of a chromosome. For each chromosome in a generation, training data are clustered by subtractive clustering algorithm with clustering parameters contained in the chromosome and a FIS is built. Both training and validation data are classified by the resulted FIS and the classification performance for each data set is evaluated. GA selects parent chromosomes to produce the next generation using a fitness function which incorporates the evaluated performances. Including classification performance of validation data in fitness function is necessary to prevent over-fitting training data by generating too many rules.

Three different metrics are used to evaluate classification performance: accuracy, sensitivity and specificity. Accuracy is defined as the ratio of the number of all correctly classified data points to the number of overall data points:

$$Ac = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

where TP (true positive) is the number of correctly classified positives, TN (true negative) is the number of correctly classified negatives, FP (false positive) is the number of negatives misclassified as positives, and FN (false negative) is the number of positives misclassified as negatives.

Accuracy gives an overall indication of classification performance; however, it does not provide an insight into classification performance for each class separately. To measure the classification performance for positive and negative classes separately, sensitivity and specificity are used respectively. Sensitivity is defined as the ratio of the number of correctly classified positives to the number of all actual positives:

$$Sn = \frac{TP}{TP + FN} \tag{5}$$

Similarly, specificity is defined as the ratio of the number of correctly classified negatives to the number of all actual negatives:

$$Sp = \frac{TN}{TN + FP} \tag{6}$$

Using these three metrics for validation and training data, the fitness function can be written as:

$$\begin{aligned} F=&W_1{*}Ac_{tr}+W_2{*}Sn_{tr}+W_3{*}Sp_{tr}+\\ &W_4{*}Ac_{va}+W_5{*}Sn_{va}+W_6{*}Sp_{va} \end{aligned} \tag{7}$$

where Ac_{tr}, Sn_{tr} and Sp_{tr} are the training accuracy, sensitivity, and specificity respectively, Ac_{av}, Sn_{av} and Sp_{av} are the validation accuracy, sensitivity, and specificity respectively, and W_1 to W_6 are their corresponding weights.

Table I. Chromosome structure for subtractive clustering optimization.

Radius1	Radius 2	……	Radius n
Squash factor	Accept ratio	Reject ratio	

The final step of proposed method is fine-tuning of input membership functions. This is done by a GA whose chromosomes contain a tuning variable for each membership function parameter. The input membership function used is Gaussian function which has two parameters. Gaussian function has a symmetric bell shape and is defined as follows:

$$G(x,\sigma,c) = e^{\frac{-(x-c)^2}{2\sigma^2}} \tag{8}$$

where c is the center of the of the bell shape and σ (standard deviation) controls the width of the bell. The chromosome structure is shown in Table II, where c_{ij} and σ_{ij} are the center and standard deviation of the jth membership function of the ith input.

ISBN: 978-1-61804-227-9

Table II. Chromosome structure for input membership function parameters optimization.

c_{11}	σ_{11}	c_{12}	σ_{12}	…	…	c_{1j}	σ_{1j}
c_{21}	σ_{21}	c_{22}	σ_{22}	…	…	c_{2j}	σ_{2j}
:	:	:	:	:	:	:	:
c_{i1}	σ_{i1}	c_{i2}	σ_{i2}	…	…	c_{ij}	σ_{ij}

V. Experimental Study

To evaluate performance of the proposed method, three different data sets are used for testing. Each data set is divided into two sets: modeling data (70%) and testing data (30%). Modeling data are used to build a FIS using the proposed method and to train the resulted FIS using a neuro-fuzzy system. The neuro-fuzzy system used is "Adaptive Network-based Fuzzy Inference System" (ANFIS). ANFIS, introduced by Jang (1993), is a method for training FIS parameters through a hybrid learning rule which combines the back-propagation gradient descent method and a least-squares method [16]. Accuracy, sensitivity and specificity of the resulted FIS are computed for each data set using testing data partition. These results are compared to results of a number of ANFIS models whose fuzzy rules are generated by fuzzy c-means (FCM) clustering method with various cluster numbers (number of rules). Testing results for each model are the average results of multiple evaluations each with different random data sampling. Number of GA population used is 40 and maximum number of generations is 100.

The three data sets used in testing are: Pima data set, BUPA data set and bladder cancer data set. All data are normalized before being used in training or testing. Pima and BUPA data sets are taken from University of California in Irvine (UCI) Machine Learning Repository [17].

A. *Pima data set*

Pima data set contains 768 records of Pima Indian patients from USA tested for diabetes. All patients are females at least 21 years old. Table III shows details of Pima data set.

Table III. Details of Pima data set.

Number of instances	768
Number of +ve/-ve	268/500
Number of attributes	9

Classification results of BUPA data set are shown in Table IV. The first eight rows are results of classification using 8 FIS's trained by ANFIS and generated by fuzzy c-means clustering with different number of clusters. Bold results indicate best results among these rows. The last row shows the results of classification using a FIS generated by the proposed method (Subtractive Clustering and Genetic Algorithm, SCGA) and trained by ANFIS.

As shown in the Table IV, the accuracy and sensitivity of the proposed method are higher than the highest accuracy and sensitivity of the eight ANFIS models. However, there are two ANFIS models whose specificities are higher than that of the proposed method. In general, the proposed method tries to find the optimal solution which has the best classification performance with respect to all the three metrics. Fig. 2 and Fig. 3 show plots of sensitivity and specificity respectively, for the eight ANFIS models. The dotted line in each figure represents the result of the proposed methods. These figures clearly show the tradeoff between sensitivity and specificity.

Table IV. Classification results for Pima data set.

Rule generation method	Number of rules	Accuracy (%)	Sensitivity (%)	Specificity (%)
FCM	2	76.30	56.60	86.60
FCM	3	**77.10**	52.20	89.30
FCM	5	76.60	57.00	86.50
FCM	10	75.80	**57.20**	85.90
FCM	15	72.40	55.00	81.50
FCM	25	70.87	38.69	86.64
FCM	40	67.39	18.54	**93.18**
FCM	50	65.83	27.17	86.83
SCGA	Auto	77.13	58.30	87.12

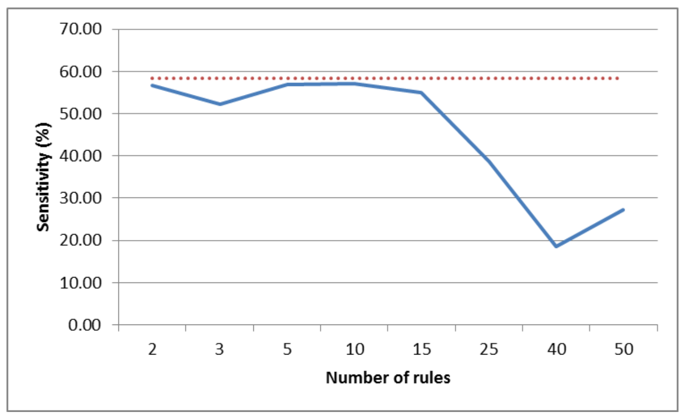


Fig.2. Sensitivity versus number of rules for the eight ANFIS models for Pima data set.

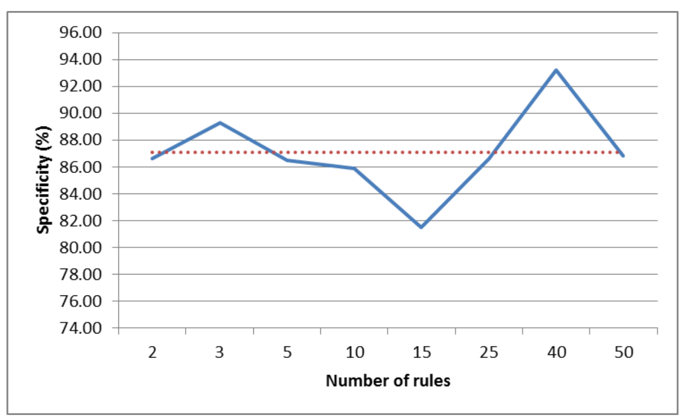


Fig.3. Specificity versus number of rules for the eight ANFIS models for Pima data set.

Fig. 4 and Fig 5 depict plots of testing and training accuracies respectively. The dotted line in each figure represents the result of the proposed methods. It is noticed that increasing the number of rules increases training accuracy; however, it reduces testing accuracy (over-fitting). The proposed method tries to avoid the problem of over-fitting by finding the optimal solution for both training and validation.

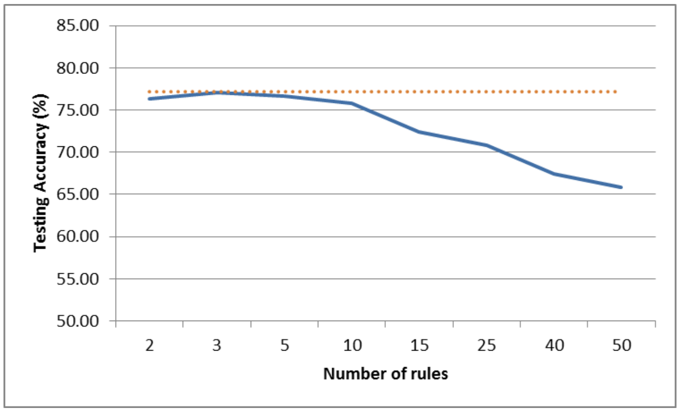


Fig.4. Testing accuracy versus number of rules for the eight ANFIS models for Pima data set.

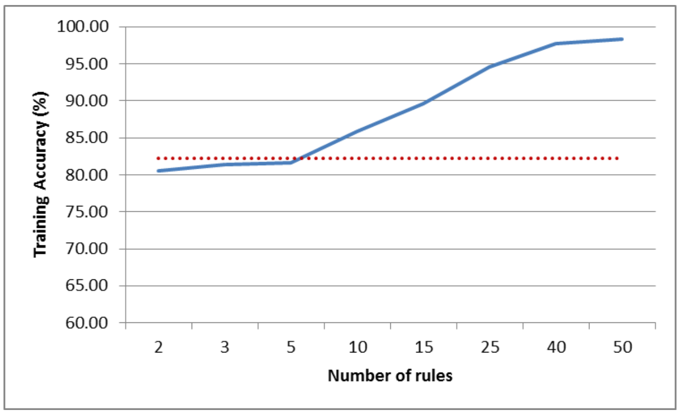


Fig.5. Training accuracy versus number of rules for the eight ANFIS models for Pima data set.

B. BUPA data set

BUPA data set contains 345 records of liver patients from USA. Details of BUPA data set are shown in Table V.

As with Pima data set, BUPA data set is used to compare classification performance of the proposed method to the performance of eight ANFIS models. The results are shown in Table VI. The proposed method has higher accuracy than all the eight models and sensitivity that is very close to the highest one. The plots of testing and training accuracies are shown in Fig. 6 and Fig. 7 respectively.

Table V. Details of BUPA data set.

Number of instances	345
Number of +ve/-ve	200/145
Number of attributes	7

Table VI. Classification results for BUPA data set.

Rule generation method	Number of rules	Accuracy (%)	Sensitivity (%)	Specificity (%)
FCM	2	70.20	79.10	58.30
FCM	3	**71.35**	79.05	59.56
FCM	5	70.40	87.30	45.80
FCM	10	64.80	69.80	58.10
FCM	15	64.20	**88.60**	24.70
FCM	25	64.81	73.66	53.56
FCM	40	63.85	64.01	**61.06**
FCM	50	65.77	80.22	44.26
SCGA	Auto	71.92	78.29	60.96

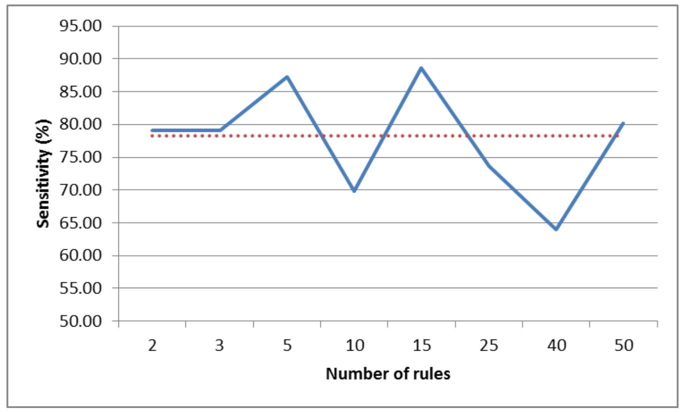


Fig.6. Sensitivity versus number of rules for the eight ANFIS models for BUPA data set.

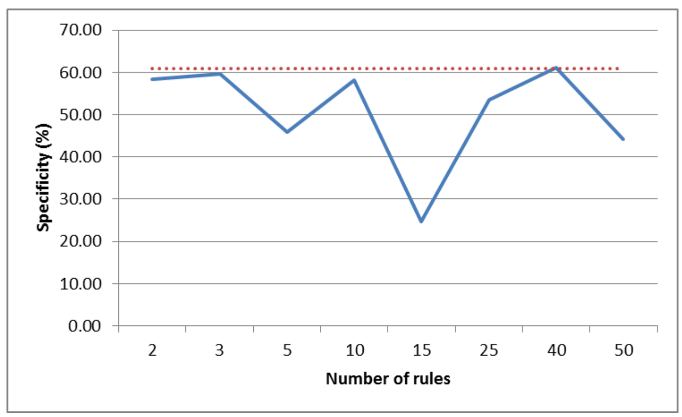


Fig.7. Specificity versus number of rules for the eight ANFIS models for BUPA data set.

C. Bladder cancer data set

Bladder cancer data set comprises of progression information of 234 patients who had undergone surgical tumor removal for bladder cancer. Table VII shows details of this data set. Results of classification testing and comparison are shown in Table VIII. The plots of testing and training accuracies are shown in Fig. 6 and Fig. 7 respectively.

Table VII. Details of bladder cancer data set.

Number of instances	234
Number of +ve/-ve	163/71
Number of attributes	8

Table VIII. Classification results for bladder cancer data set.

Rule generation method	Number of rules	Accuracy (%)	Sensitivity (%)	Specificity (%)
FCM	2	71.70	20.00	94.40
FCM	3	**75.90**	16.20	98.90
FCM	5	70.30	**26.20**	90.10
FCM	10	74.50	26.00	94.80
FCM	15	64.10	24.50	89.70
FCM	25	62.80	8.25	97.60
FCM	40	66.20	8.00	**100.00**
FCM	50	73.10	10.60	97.60
SCGA	Auto	77.93	14.44	**99.00**

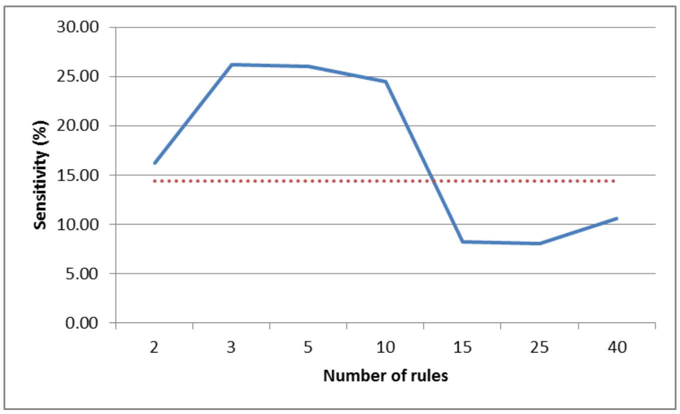


Fig.8. Sensitivity versus number of rules for the eight ANFIS models for bladder cancer data set.

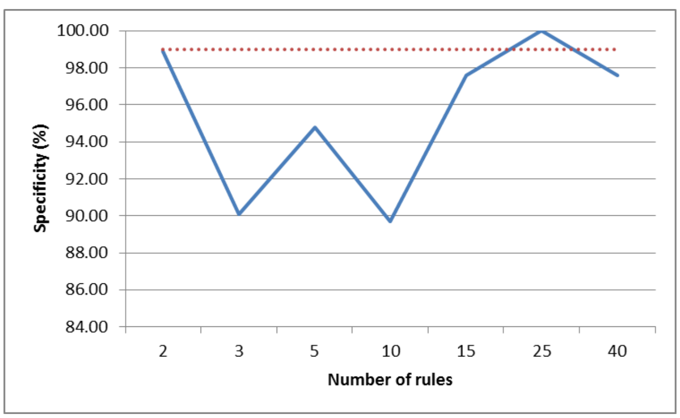


Fig.9. Specificity versus number of rules for the eight ANFIS models for bladder cancer data set.

VI. Conclusion

In this paper, we proposed a method based on subtractive clustering and genetic algorithm to automatically generate fuzzy rules for data classification. The proposed method searches for the FIS structure and the number of rules that best compromise between classification performance of training and validation. Experimental results show that the proposed method has better accuracy than one of the widely used clustering method (fuzzy c-means) and a well compromised sensitivity and specificity. Results also show that FIS generated by the proposed method avoids data over-fitting.

References

[1] L. Zadeh, “Fuzzy sets,” Inf. Control, vol. 353, pp. 338–353, 1965.
[2] M. J. Er and Y. Zhou, “Automatic generation of fuzzy inference systems via unsupervised learning,” Neural Netw., vol. 21, no. 10, pp. 1556–66, Dec. 2008.
[3] S. L. Chiu, “Extracting fuzzy rules from data for function approximation and pattern classification,” Fuzzy Information Engineering: A Guided Tour of Applications, pp. 1–10, 1997.
[4] A. F. Gómez-Skarmeta, M. Delgado, and M. A. Vila, “About the use of fuzzy clustering techniques for fuzzy model identification,” Fuzzy Sets Syst., vol. 106, no. 2, pp. 179–188, Sep. 1999.
[5] P. P. Angelov and R. A. Buswell, “Automatic generation of fuzzy rule-based models from data by genetic algorithms,” Inf. Sci. (Ny)., vol. 150, pp. 17–31, 2003.
[6] R. P. Prado, S. García-Galán, J. E. Exposito, and A. J. Yuste “Knowledge acquisition in fuzzy-rule-based systems with particle-swarm optimization,” Fuzzy Syst. IEEE Trans., vol. 18, no. 6, pp. 1083–1097, 2010.
[7] H. Zhang, B. Zhang, and F. Wang, “Automatic Fuzzy Rules Generation Using Fuzzy Genetic Algorithm,” 2009 Sixth Int. Conf. Fuzzy Syst. Knowl. Discov., pp. 107–112, 2009.
[8] L. Xiao and G. Liu, “Automatic fuzzy rule extraction based on fuzzy neural network,” Adv. Neural Networks–ISNN 2005, no. 104262, Springer Berlin Heidelberg, pp. 688–693, 2005.
[9] M. Sikora, “Fuzzy Rules Generation Method for Classification Problems Using Rough Sets and Induction of Decision Rules Based Upon Rough Set,” in Lecture Notes in Computer Science Volume, 2005, pp. 383–391.
[10] D. Wang, T. S. Dillon and E. J. Chang, “A data mining approach for fuzzy classification rule generation,” Proc. Jt. 9th IFSA World Congr. 20th NAFIPS Int. Conf. (Cat. No. 01TH8569), vol. 5, no. C, pp. 2960–2964, 2001.
[11] E. H. Mamdani and S. Assilian, “An experiment in linguistic synthesis with a fuzzy logic controller,” Int. J. Man. Mach. Stud., 1975.
[12] T. Takagi and M. Sugeno, “Fuzzy identification of systems and its applications to modeling and control,” IEEE Trans. Syst. Man. Cybern., vol. SMC-15, no. 1, pp. 116–132, Jan. 1985.
[13] C. Z. Janikow and Z. Michalewicz, “An experimental comparison of binary and floating point representations in genetic algorithms.,” ICGA, 1991.
[14] D. E. Goldberg, “Real-coded genetic algorithms, virtual alphabets, and blocking,” Urbana, pp. 1–21, 1990.
[15] C. R. Houck, J. A. Joines, and M. G. Kay, “A genetic algorithm for function optimization: a Matlab implementation,” NCSU-IE TR, no. 919, pp. 1–14, 1995.
[16] J. Jang, “ANFIS: adaptive-network-based fuzzy inference system,” Syst. Man Cybern. IEEE Trans. …, vol. 23, no. 3, 1993.
[17] A. Frank and A. Asuncion, UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science, 2010.

An application of Neural Networks for Prediction of Surface Texture Parameters in Turning

S. Karagiannis, P. Stavropoulos and J. Kechagias

Abstract— Surface roughness, an indicator of surface quality is one of the most specified customer requirements in a machining process. Mastering of surface quality issues helps avoiding failure, enhances component integrity, and reduces overall costs. Copper alloy (GC-CuSn12) surface quality, achieved in turning, constitutes the subject of the current research study. Test specimens in the form of near-to-net-shape bars and a titanium nitride coated cemented carbide (T30) cutting tool were used. The independent variables considered were the tool nose radius (r), feed rate (f), cutting speed (V), and depth of cut (a). Process performance is estimated using the statistical surface texture parameters Rα, Ry, and Rz. To predict the surface roughness, an artificial feed forward back propagation neural network (ANN) model was designed for the data obtained.

Keywords— Neural Networks, Modelling, Turning, Surface Texture Parameters.

I. Introduction

Copper-based alloys are used in the mass production of electrical components and water pipe fittings. They are usually machined using high speed CNC machines, which are mostly very high speed lathes fed with brass wire of a relatively small diameter, so that the maximum speed is limited to 140-220m/min, although the tooling is capable of a good performance at much higher speeds. When copper alloys are machined, very high forces act on the tool, particularly at low cutting speeds. This is due to the large contact area on the rake face resulting in a small shear plane angle and thick chips [1], contributing in the fact that copper is cosidered as one of the most difficult materials to machine. When feed rate is decreased or the cutting speed is increased the cutting forces are decreased and the surface finish is improved.

S. Karagiannis is with Department of Mechanical Engineering, Technological Educational Institute of Western Macedonia, Koila Kozanis, Greece,(email: skaragiannis@teikoz.gr)

P. Stavropoulos is with Laboratory for Manufacturing Systems and Automation, Department of Mechanical Engineering and Aeronautics, University of Patras, Patras 265 00, Greece, (email: pstavr@lms.mech.upatras.gr)

J. Kechagias is with Department of Mechanical Engineering, Technological Educational Institute of Thessaly, Larissa, Greece (corresponding author: phone: 0030 2410684322, fax: 0030 2410684305, email: jkechag@teilar.gr)

Surface properties dominate the quality of the finished component, since they influence features like dimensional accuracy; tribological issues such as the friction coefficient and wear; post processing requirements; appearance and cost. Surface roughness or texture constitutes a measure for achieving finer surface irregularities in the finished product, while three components –i.e., roughness, waviness, and form– are required for its determination [2].

A number of studies –investigating the relation of cutting forces, tool wear, chip morphology, accuracy issues, and dynamic behaviour during turning with the produced surface quality– are reported in literature. A study of the effects of different process parameters: tool radius (r), feed rate (f), cutting speed (V), and depth of cut (a) in turning of a copper alloy (GC-CuSn12), on the surface texture parameters Ra, Rz, and Ry is attempted in the current work, using the Taguchi methodology and Neural Networks modelling.

Thus, an $L_9(3^4)$ orthogonal matrix experiment was conducted [3]. A matrix experiment consists of a set of experiments where the settings of several process parameters to be studied are changed from one experiment to another in a combinatory way. Experimental results are used in order to train a feed forward back propagation neural network (FFBP-NN) in order to predict surface texture parameters in turning of near-to-net shape parts of copper alloy. Using FFBP-NN in combination with orthogonal matrix experiment, an easy way modeling could be achieved, and applied on experimental region in order to predict surface texture parameters.

II. Experimental Setup

The material used for cutting is specified as GC-CuSn12. It is a copper alloy containing 84 to 85% Cu, 11 to 14% Zn, under 1% Pb, less than 2% Ni, and finally under 0.2% Sb. The machine used for the experiments was a Cortini F100 CNC machine lathe (3.7kW) equipped with a GE Fanuc Series O-T control unit. The test specimens were in the form of bars, 32mm in diameter and 80mm in length for near-to-net-shape machining. Tailstock was not used (Fig. 1). The cutting tools were titanium nitride screw-on positive inserts, CCMT 09T30, with a 0.4 and 0.8mm tool nose radii, accordingly (Fig. 2).

Surface roughness is a widely used index characterising a product's quality, and is measured off-line –when the

ISBN: 978-1-61804-227-9

component is already machined. The surface texture parameters measured during this study are: the average surface roughness (also known as centre line average – CLA), $R\alpha$; average maximum peak to valley height of the profile, R_z; and maximum peak to valley height; R_y; all measured in μm. Measurements are being conducted using the Mitutoyo, surftest RJ-210 tester.

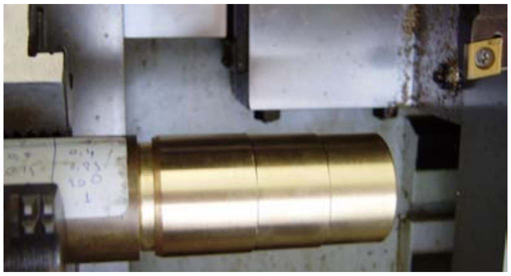


Fig. 1: Cortini F100 CNC machine lathe


Fig. 2: Machined specimens and inserts

The CLA Ra (Fig. 3) can be obtained by taking the arithmetic mean of the absolute values of 1150 different positional deviations over a 4 mm standard length with a cut-off at 0.8 mm according to the relation,

$$R_a = \frac{1}{1150}\sum_{1=1}^{1150}|y_i| \quad (1)$$

The average maximum peak to valley height of the profile (Rz) is defined according to the relation,

$$R_z = \frac{1}{5}\left(\sum_{i=1}^{5}\left|y_{p_i}\right| + \sum_{i=1}^{5}\left|y_{v_i}\right|\right) \quad (2)$$

where Ypi are the five tallest peaks, and Yvi, the five lowest valleys within the sample considered (Fig. 3). Finally, the maximum peak to valley distance of the filtered profile (Ry) over an evaluation length sensitive to large deviations from the mean line and scratches is defined according to the relation,

$$R_y = R_p + R_v \quad (3)$$

where Rp and Rv are the absolute values of the maximum peak and maximum valley within the measured standard length (Fig. 3). A four parameter design was performed as shown in Table 1. Note that Level 1 and level 3 for the parameter (r) assign the same value. This is not an obstacle for the methodology followed.

The Taguchi design method is a simple and robust technique for process parameters optimisation. The method involves the damping (reduction) of variation in a manufacturing process through robust design of experiments. Taguchi's emphasis on minimising deviation from target, led him to develop measures of the process output that incorporate both the location of the output as well as its variation.

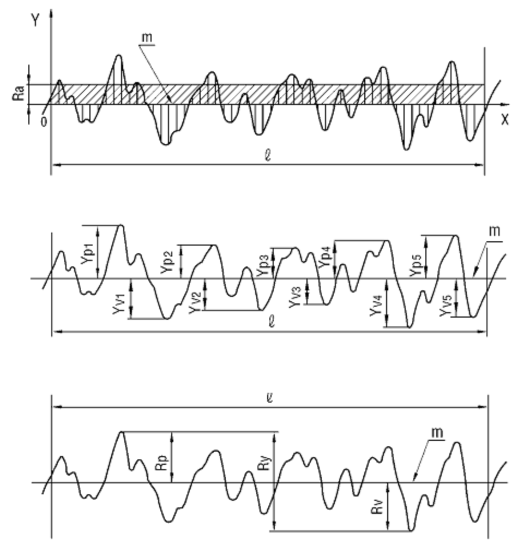


Fig. 3: Surface texture parameters

Table 1: Parameter design.

		Levels		
No	Process Parameters	1	2	3
1	Tool Radius (r, mm)	0.4	0.8	0.4
2	Feed Rate -(f, mm/rev)	0.05	0.15	0.25
3	Cutting Speed (V, m/min)	100	150	200
4	Depth of cut (a, mm)	0.2	0.6	1

These measures are called *signal-to-noise* ratios. The signal-to-noise ratio provides a measure of the impact of noise factors on performance.

Table 2: Orthogonal array $L_9(3^4)$.

	Column			
No Exp	1	2	3	4
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3	1	3	2
8	3	2	1	3
9	3	3	2	1

ISBN: 978-1-61804-227-9

Calculation of the S/N ratio depends on the *experimental objective* according to which the experiment is conducted –i.e., bigger-the-better, smaller-the-better, and nominal-is-best with corresponding calculation formulae [5]. The standard ($L_9(3^4)$) orthogonal matrix experiment was used (Table 2). Columns 1, 2, 3, and 4 are assigned to tool radius (r), feed rate (f), cutting speed (V) and depth of cut (a), respectively.

III. Experimental Results

The Taguchi design method is a simple and robust technique for optimizing the process parameters. In this method, main parameters, which are assumed to have an influence on process results, are located at different rows in a designed orthogonal array. With such an arrangement randomized experiments can be conducted. In general, signal to noise (S/N) ratio (n, dB) represents quality characteristics for the observed data in the Taguchi design of experiments. In the case of surface roughness amplitude [4][6][7], lower values are desirable. These S/N ratios in the Taguchi method are called as the smaller-the-better characteristics and are defined as follows:

$$\eta = -10\log_{10}\left[\frac{1}{n}\sum_{i=1}^{n} y_i^2\right] \qquad (4)$$

where y_i is the observed data at the i^{th} trial and n is the number of trials. From the S/N ratio, the effective parameters having an influence on process results can be obtained and the optimal sets of process parameters can be determined. Based on Robust design, the standard orthogonal array ($L_9(3^4)$) has been selected in order to perform the matrix experiment (Table 3). Three levels for each factor were selected (Table 1). Following the ($L_9(3^4)$) orthogonal array nine experiments were performed with each experiment producing a test part which was tested for R_a, Rz, and Ry all measured in μm.

Table 3: Matrix experiment

Ex. No.	r	f	V	a	R_a	R_z	Ry
1	0.4	0.05	100	0.2	1.713	9.238	11.154
2	0.4	0.15	150	0.6	1.384	9.020	11.001
3	0.4	0.25	200	1	4.576	19.290	22.488
4	0.8	0.05	150	1	1.278	7.972	9.947
5	0.8	0.15	200	0.2	1.674	10.151	12.024
6	0.8	0.25	100	0.6	2.352	13.688	16.055
7	0.4	0.05	200	0.6	1.549	9.208	10.663
8	0.4	0.15	100	1	1.995	13.862	16.604
9	0.4	0.25	150	0.2	5.149	22.005	24.761
Mean (m)					2.407	12.714	14.966

IV. Neural Network Architecture

Aiming in the prediction of the produced surface roughness parameters (Ra, Rz, and Ry) during longitudinal turning of a Cooper alloy, a NN model has been developed. The four (4) factors studied were used as input parameters of the NN model. Previous studies [8] indicate that by using Taguchi's DoE methods, a structured method of NN parameter-setting can be implemented, which identify NN and training parameter settings resulting in enhanced NN performance. Training samples are presented to the NN during training, and the network is adjusted according to its error. The nine (9) experimental data samples (Table 3), were separated into three groups, namely the training, the validation and the testing samples. Training samples are presented to the network during training and the network is adjusted according to its error. Validation samples are used to measure network generalization and to halt training when generalization stops improving. Testing samples have no effect on training and so provide an independent measure of network performance during and after training (confirmation runs).

In general, a standard procedure for calculating the proper number of hidden layers and neurons does not exist. For complicated systems the theorem of Kolmogorov or the Widrow rule can be used for calculating the number of hidden neurons [9]. In this work, the feed-forward with back-propagation learning (FFBP) architecture has been selected to analyze the surface texture parameters. These types of networks have an input layer of X inputs, one or more hidden layers with several neurons and an output layer of Y outputs. In the selected ANN, the transfer function of the hidden layer is hyperbolic tangent sigmoid, while for the output layer a linear transfer function was used. The input vector consists of the four process parameters of Table 3. The output layer consists of the performance measures, namely the Ra, Rz, and Ry surface texture parameters. According to ANN theory FFBP-NNs with one hidden layer are appropriate to model each mapping between process parameters and performance measures in engineering problems [10].

In the present work, five trials using FFBP-NNs with one hidden layer were tested having 10, 11, 12, 13 and 14 neurons each; see Figure 5. This one that has 13 neurons on the hidden layer gave the best performance as indicated from the results tabulated in Table 4.

The one-hidden-layer 13-neurons FFBB-NN was trained using the Levenberg-Marquardt algorithm (TRAINLM) and mean square error (MSE) used as objective function. The data used were randomly divided into three subsets, namely the training, the validation and the testing samples.

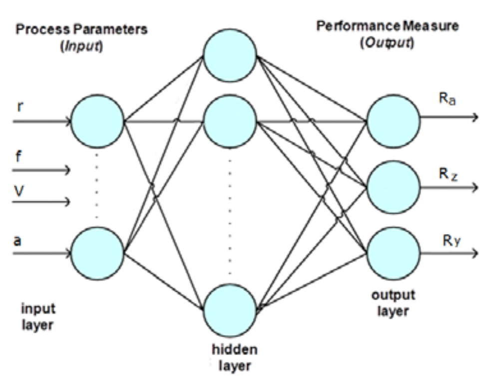


Fig. 4: The selected ANN architecture (feed-forward with back-propagation learning).

Table 4. Best performance of ANN architecture.

	ANN Architecture				
	4x10x3	**4x11x3**	**4x12x3**	**4x13x3**	**4x14x3**
Training	1	1	1	1	1
Validation	0.8963	0.8440	0.8719	0.9544	0.7224
Test	0.1265	0.8557	0.8292	0.99.06	0.8492
All	0.8547	0.9333	0.9541	0.998-	0.9246
Best val. perf.	9.88	11.64	11.17	**2.91**	14.82
epoch	1	2	4	1	1

Back-propagation ANNs are prone to the overtraining problem that could limit their generalization capability [8]. Overtraining usually occurs in ANNs with a lot of degrees of freedom [10] and after a number of learning loops, in which the performance of the training data set increases, while the performance of the validation data set decreases. Mean Squared Error (MSE) is the average squared difference between network output values and target values. Lower values are better. Zero means no error. The best validation performance is equal to 2.91 at epoch 1; see Figure 5.

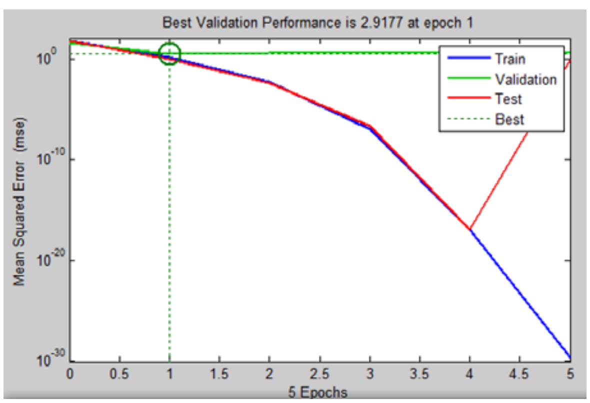


Fig. 5: The selected ANN architecture (feed-forward with back-propagation learning).

Another performance measure for the network efficiency is the regression (R); see Figure 6. Regression values measure the correlation between output values and targets. The acquired results show a good correlation between output values and targets during training (R=1), validation (R=0.9544), and testing procedure (R=0.9906).

The trained ANN model can be used for the optimization of the cutting parameters during longitudinal turning of a cooper alloy.

This can be done by testing the behaviour of the response variable (Ra, Rz, and Ry) under different variations in the values of tool radius (r), feed rate (f), cutting speed (V), and depth of cut (a) (Fig. 7).

V. Conclusions

The surface texture parameters (R_a, R_z, and R_y) of copper alloy near-to-net-shape parts during turning was measured according to a matrix experiment. The results were used to train a feed forward back propagation neural network with a topology of 4X13X3 neurons. The proposed NN can be used to predict the surface texture parameters as well as to optimize the process according to each one of the surface texture parameters. As a future work Authors plan to improve the performance of FFBP-NN incorporating more experiments as well as investigate the performance of alternatives training algorithms. In addition a comparison among other approaches such as regression and additive modeling will be performed. Using the extracted NN the surface response of R_a, R_z, and R_y can be drawn and the effects of process parameters be estimated inside the experimental region in which the designed experiment is conducted. This methodology could be easily applied to different materials and initial conditions for optimization of other material removal processes.

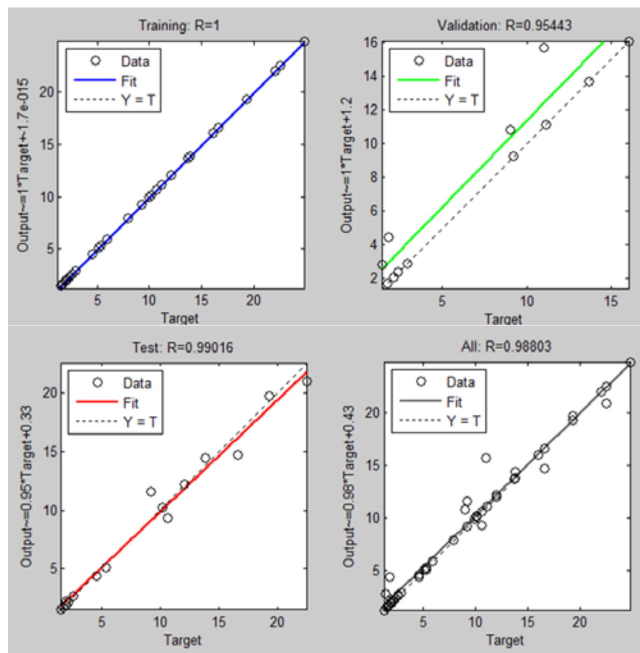


Fig. 6: Regression plots

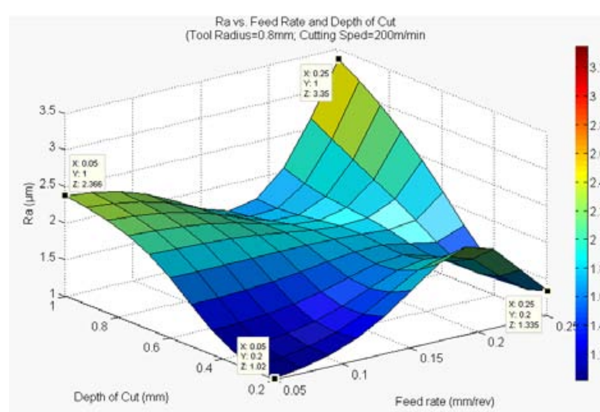


Fig. 7: CLA Ra according feed rate and depth of cut (r=0.8mm, V=200m/min)

ACKNOWLEDGEMENTS

This research is implemented through the Operational Program 'Education and Lifelong Learning' and is co- financed by the European Union (European Social Fund) and Greek national funds.

REFERENCES

[1] Trent, E.M., Wright, P.K. (2000). Metal Cutting. Butterworth-Heinemann, Boston.

[2] Karagiannis S., Stavropoulos P., Ziogas C., Kechagias J., (2013), "Prediction of surface roughness magnitude in computer numerical controlled end milling processes using neural networks, by considering a set of influence parameters: An aluminium alloy 5083 case study", Proc IMechE Part B: J Engineering Manufacture, DOI: 10.1177/0954405413498582

[3] Phadke, M.S., (1989). Quality Engineering using Robust Design. Prentice-Hall, Englewood Cliffs, NJ.

[4] Kechagias, J., (2007). "An experimental investigation of the surface roughness of parts produced by LOM process". Rapid Prototyping J. Vol. 13 No 1, p. 17-22.

[5] Montgomery, D., (2008), "Design and Analysis of Experiments", 7th Edition, John Wiley & Sons.

[6] Petropoulos, G., Mata, F., Davim, J. P., (2008). "Statistical study of surface roughness in turning of peek composites". Materials & Design, Vol. 29, No. 1, p. 218-223.

[7] Tsao, C. C., (2009). "Grey–Taguchi method to optimize the milling parameters of aluminum alloy". The International Journal of Advanced Manufacturing Technology. Vol. 40, p. 41–48.

[8] Sukthomaya, W., Tannock, J., (2005), "The optimisation of neural network parameters using Taguchi's design of experiments approach: an application in manufacturing process modelling", Neural Computing and Applications, Volume 14, Issue 4, pp 337-344

[9] Tzafestas, S.G., et al., (1996). "On the overtraining phenomenon of backpropagation NNs.", Mathematics and Computers in Simulation, 40, 507-521.

[10] Lin, C.T., Lee, G.C.S., (1996), "Neural fuzzy systems- A neuro-fuzzy synergism to intelligent systems". Prentice Hall PTR, p. 205-211.

[11] Prechelt, L., (1998), " Automatic early stopping using cross validation: quantifying the criteria.". Neural Networks, 11(4), 761–767.

Evolutionary Computing Optimization based Extreme Learning Machine for Pulmonary Data Classification

P.V. Pramila and V. Mahesh

Abstract— Chronic obstructive pulmonary disease (COPD) is a respiratory syndrome associated with a progressive, non-reversible limitation to airflow. The survey of Global Strategy for the Prevention and Control of Non communicable Diseases reveals that almost 90% deaths occur due to COPD in low-and middle-income countries surpassing the cardiovascular disease, causing a huge health and economic burden. In this work an attempt has been made to interpret and to classify the obstructive pulmonary disorder with the support of computational intelligence techniques. Neural Computing and Genetic Computing is employed to classify and the performance of the networks are analyzed. The results show that the performance of Extreme Learning Machine (ELM), achieved a generalization of 91.03% of accuracy in 0.0019secs while the global search evolutionary based optimization technique achieved an accuracy of 100% with a much compact network. Results indicate that the performance of Evolutionary Extreme Learning Machine (EELM) significantly improve the generalization ability of ELM with a less complexity, compact network and provide assistance to clinicians to diagnose and treat pulmonary disorders in incomplete data sets.

***Keywords*— COPD, Spirometry, Classifier, Extreme learning machine, Flow- volume loop.**

I. Introduction

The respiratory system is one of the most important elements in maintaining vital functions of the human body. The large complexity of its mechanics is a well-known issue [1,2]. A survey report by the Global Strategy for the Prevention and Control of Noncommunicable Diseases (2010) reveals that of the 57 million deaths that occurred globally in 2008, 36 million – almost two thirds – were due to Non Communicable Disease(NCDs), comprising mainly cardiovascular diseases, cancers, diabetes and chronic lung diseases[3]. A regular, healthy functioning lung depends on many mechanical and physiological factors and there are pathologies which disturb the symmetry in this natural organisation. Obstructive and Restrictive diseases are two major abnormality conditions of the lungs. In obstructive lung condition, the airways are narrowed, usually causing an increase in the time it takes to empty the lungs.

P.V. Pramila, Associate Prof., Dept. of ECE, Dhanalakshmi College of Engineering, Chennai ,India.

Dr. V. Mahesh , Associate Prof., Dept. of BME, SSN College of Engineering ,Chennai, India Phone:+91-44-27469700; fax:+91-44-27469772; e-mail: maheshv@ssn.edu.in.

The obstructive condition is characterized by progressive airflow obstruction of the peripheral airways associated with lung inflammation, emphysema and mucus hyper secretion [4]. Over 80% of cardiovascular and diabetes deaths, and almost 90% of deaths from chronic obstructive pulmonary disease, occur in low- and middle-income countries[3]. Chronic obstructive pulmonary disease (COPD) is a respiratory syndrome associated with a progressive, non-reversible limitation to airflow and abnormal inflammatory responses involving the smaller airways [5,6]. It is observed that chronic obstructive pulmonary disease (COPD) affecting people worldwide presents a huge health and economic burden.

Current diagnosis of COPD involves an assessment of smoking and occupational history as well as recording of symptoms such as cough, sputum and dyspnoea. However, due to poor public awareness of this disease and the overlap of symptoms with other co-morbidities, many patients are not diagnosed until later when the disease has made a significant impact on their quality of life [7, 8]. In COPD patients, compared with healthy elderly subjects, walking time and frequency of postural changes is significantly decreased [9]. Sleep quality is often poor in patients with COPD, and patients frequently report difficulty in falling asleep, nocturnal awakening, and insomnia. [10-12]. In recent years, increasing attention has been paid to the role of co-morbidities in patients with chronic obstructive pulmonary disease (COPD). The risk ratio of developing Heart Failure among COPD patients is 4.5 times higher than that of control individuals without the disease, after adjusting for age and other cardiovascular risk factors [13,14]. COPD is a well known independent risk factor for the development of post operative pulmonary complications after thoracic or non thoracic surgery.[15,16]

Spirometry is the standardized and an objective way of measuring airflow obstruction for the diagnosis and management of chronic obstructive pulmonary disease (COPD).[17,18]. Air flow obstruction, either acute or chronic, will reduce the forced expiration volume in 1sec (FEV1) by increasing the airway resistance to expiratory flow. Due to premature closing of the airways on expiration, the Forced Vital Capacity (FVC) will also decrease but not as much as the FEV1. The ratio of two significant parameters considered for COPD diagnosis (FEV1/FVC) will decrease and the flow/volume graph has a concave appearance (Fig. 1). The volume/time graph will show a line that rises slowly to reach its highest point, completing the full expiratory manoeuvre (Fig. 1).

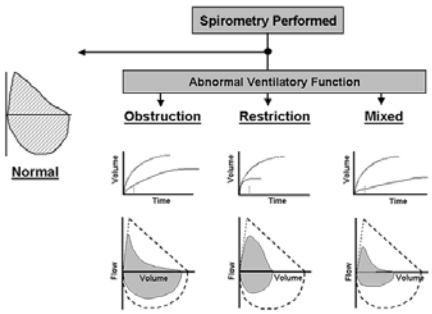


Fig. 1 Flow volume loop for normal, obstructive , restrictive and mixed pulmonary disorders

COPD also termed as irreversible lung function is reflected by the reduction in FEV1 (Forced Expiratory Volume in 1 sec) and reduced ratio of FEV1/FVC (Forced vital capacity). Spirometry is also used to assess the severity of disease. Spirometer can distinguish asthma from COPD on the basis of a significant increase of 12 percent improvement in the FEV_1 measured after the use of bronchodilators and is thus a good prognostic indicator[19,20].

Despite spirometry has the distinct advantage of being a reproducible measurement of lung function that is superior to peak expiratory flow because of greater reliability and specificity, symptom based and clinical diagnosis contributes to misdiagnosis and mistreatment [21,22]. International studies have examined the prevalence of COPD misdiagnosis due to lack of Spirometry [23]. A combination of monitoring systems on the use of spirometry in COPD, more education on the importance of spirometry in COPD management, and assistance in interpretation of spirometry results may bring about improvements [24].

The objective of this work is to provide assistance to the clinicians in the interpretation and classification of Spirometry results with the support computational intelligence techniques. In this paper well known Artificial Intelligence techniques viz. Neural Computing and Genetic Computing or Evolutionary Computing is employed to classify and the performance of these networks are analyzed.

II. MATERIALS AND METHODS:

2.1 DATA ACQUISITION:

In this work, a total of 70 subjects were considered and data was acquired using InspireX Spirometer. The age, height and weight of each subject were recorded prior to the test. The subjects were advised to inhale to their full lung capacity and then exhale forcibly and the data was recorded. Spirometry with flow volume loops were obtained and each participant completed up to three efforts to measure the FEV_1 and FVC volumes according to the American Thoracic Society (ATS) standards[6].

The vital parameters in determination of pulmonary obstructive disorder are the Forced Vital Capacity (FVC) and Forced Expiratory Volume in 1sec($FEV_{1)}$. FVC is the change in lung volume from full inspiration to total lung capacity. This parameter determines how hard a subject can breathe out during a forced expiration. FEV_1 is forced expiration volume in 1sec during forced exhalation. The other lung volume values measured by a spiometer are the ratio of FVC to FEV_1, Peak expiratory flow(PEFR) and Forced expiratory flow at 75%($FEF_{25\text{-}75}$).The input data derived from the flow-volume loop are normalized between 1 and -1. The behaviour and performance of the ELM and the EELM learning algorithms with changes in initial parameters were analyzed.

2.2. EXTREME LEARNING MACHINE

Artificial neural networks analogous to human brain is coordinated to the intelligence and behaviour of human being. These computerised data modelling systems possessing properties of human brain are capable of machine learning and pattern recognition like humans. In machine learning, classification is the problem of identifying which of a set of categories (or classes) a new observation belongs to, on the basis of a number of observed attributes related to that object. The performance of the single layer feed forward neural network is evaluated for two activation function – Sigmoid and Gaussian. The choice of activation functions were from the literature [4], [24] in which the authors have established in their work that a Radial Basis Function Neural Networks (RBFNN) model can classify an obstructive abnormality from the normal subjects and sigmoid models showed a goodness of fit of $R^2 > 0.92$.

ELM is a learning algorithm for Single hidden Layer Feed forward Neural networks (SLFNs) and has the advantages of very fast learning speed than the traditional gradient-based learning algorithm, high accuracy [25], [26], [27]. In general, a multi-category classification problem can be stated in the following manner [27]. Suppose we have N observation samples $\{X_i, Y_i\}$, where $X_i= [X_{i1},\ldots\ldots,X_{in}]$ €R_n s an n-dimensional feature of the sample i and $Y_i = [Y_{i1},Y_{i2}\ldots\ldots Y_{ic}]$

€ R_c is its coded class label. If the sample X_i is assigned to the class label C_k then k^{th} element of Y_i is 1 i.e. (Y_{ik}=1) and other element are -1. Here, we assume that the samples belong to C distinct classes. The initial input weights and hidden biases of ELM are randomly chosen, and then the output weights are analytically determined by using Moore-Penrose (MP) generalized inverse.

The following are the steps involved in the ELM algorithm:

i) For a given training samples (X_i,Yi), select the appropriate activation function G (*) and the number of hidden neurons H.

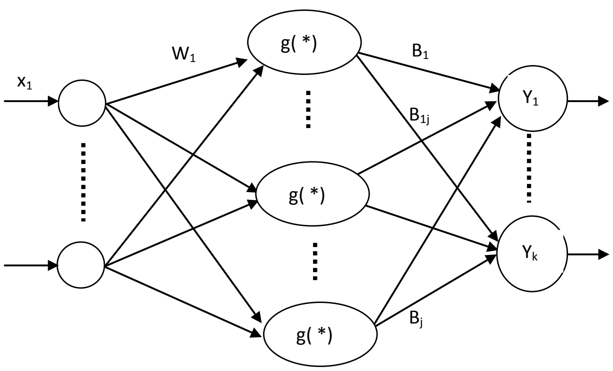


Fig.2 Single layer feed forward network

ii) For a given number of hidden neurons, it is assumed that the input weights W and bias B of hidden neurons are selected randomly. Let W be H × n input weights, B be H ×1 bias of hidden neurons and V be C × H output weights. Assuming that the predicated output $\hat{Y}$ is equal to the coded labels Y_i, the output weights are estimated analytically as,

$$y_{ik} = \sum_{j=1}^{H} V_{kj}\, G_j\,(W, B, X_j) \quad k = 1,2 \ldots C \tag{1.1}$$

In matrix form,

$$\hat{Y} = VY_H^{\dagger} \tag{1.2}$$

where $Y_H^{\dagger}$ is the Moore-Penrose generalized pseudo-inverse of the hidden layer output matrix.

Although ELM is fast and presents good generalization performance, as the output weights are computed based on the prefixed input weights and hidden biases using the Moore-Penrose (MP) generalized inverse, there may exist a set of non-optimal input weights and hidden biases [28]. Neural networks should evolve the best neural network architecture and therefore, it claims that the performance of SLFNs can be improved if the input weights and biases of hidden units are chosen properly.

2.3 EVOLUTIONARY EXTREME LEARNING MACHINE

To optimize the network, a hybrid approach called Evolutionary Extreme Learning Machine (EELM) [29],[30],[31] is employed. It utilizes the advantages of both ELM and the Evolutionary algorithm [31] i.e. the input weights and biases are determined by using the DE process and the output weights are analytically determined by using Moore-Penrose (MP) generalized inverse. The algorithm is as follows:

i). Generate the initial generation composed of parameter vectors { $\boldsymbol{\theta}_{i,0}$ | i=1, 2 ,…, NP } as the population.

ii) At each generation G, we do:

a) Mutation: the mutant vector is generated by

$$v_{j,i,G+1} = x_{r1,G} + F\left(x_{r2,G} - x_{r3,G}\right) \tag{2}$$

where r_1, r_2, and r_3 are different random indices, and F is a constant factor used to control the amplification of the differential variation and takes a value between 0 and 2.

b) Crossover: The trial vector $u_{i,G+1}$ is developed from the elements of the target vector, $x_{i,G,}$ and the elements of the donor vector, $v_{j,i,G+1}$

$$u_{j,i,g+1} = \begin{cases} v_{j,i,G+1} & \text{if } rand_{j,i} \le CR \text{ or } j = I_{rand} \\ x_{j,i,G} & \text{if } rand_{j,i} > CR \text{ or } j \ne I_{rand} \end{cases} \tag{3}$$

c) Calculate the output weight using (1.1)

d) Selection: Determination of a new population, $u_{j,i,g+1}$ using (4). In addition to the root mean square error, the norm of the output weight ||A|| is also used as a criteria to reinforce the selection.

$$x_{i,G+1} = \begin{cases} u_{i,G} & \text{if } f(x_{i,G}) - f(u_{i,G+1}) > \epsilon f(x_{i,G}) \\ u_{i,G} & \text{if } \left|f(x_{i,G}) - f(u_{i,G+1})\right| < \epsilon f(x_{i,G}) \\ & \text{and } \|A^{u_{i,G}}\| < \|A^{x_{i,G}}\| \\ x_{i,G} & \text{otherwise} \end{cases} \tag{4}$$

iii) Iteration of the above process once the new population is generated, until the goal is met or a preset maximum number of learning iterations is reached.

III. RESULTS AND DISCUSSION:

Statistical analysis of mean and standard deviation of the input parameters is shown in Table 1. The mean and standard deviation values show a significant variation between the normal and abnormal subjects.

ISBN: 978-1-61804-227-9

TABLE I STATISTICAL ANALYSIS OF INPUT PARAMETERS OF NORMAL AND ABNORMAL SUBJECTS

Parameter	Normal	Abnormal
	Mean± SD	***Mean± SD***
FVC	3.24 ± 0.57	2.19 ± 0.85
FEV1	2.71 ± 0.51	1.25 ± 0.58
FEV1/FVC	84 ± 0.08	57 ± 0.2
PEFR	2.13 ± 3.02	2.95 ± 2.13
FEF	1.61 ± 0.87	0.68 ± 0.92

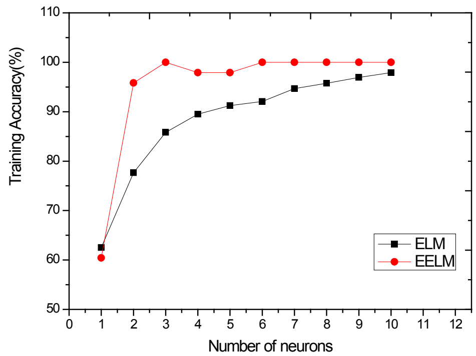


Fig. 3 The Training Accuracy of the ELM and EELM classifiers on a wide range of hidden neurons and *Radbas activation* function

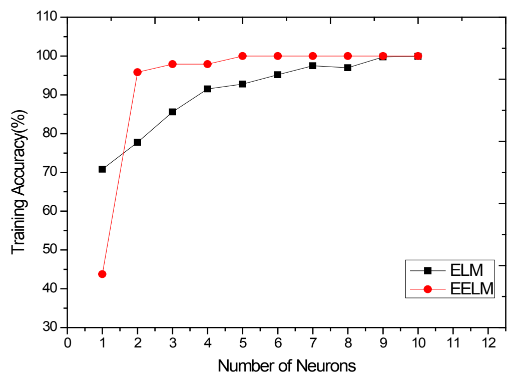


Fig. 5 The Training Accuracy of the ELM and EELM classifiers on a wide range of hidden neurons and *Sigmoid* activation function

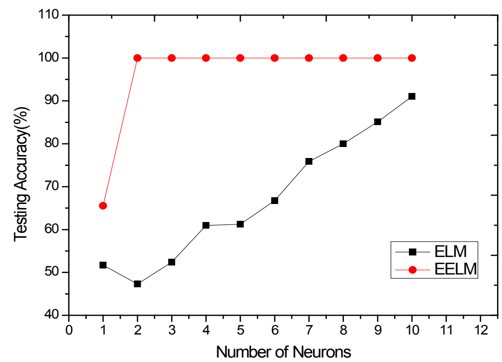


Fig. 6 The Testing Accuracy of the ELM and EELM classifiers on a wide range of hidden neurons and *Sigmoid* activation function

TABLE II MEAN TESTING EFFICIENCY AND PROCESSING TIME FOR ALGORITHMS – ELM, EELM WITH RADBAS ACTIVATION FUNCTION

Hidden Neurons	**Training Time(secs)**		**Training Accuracy (%)**		**Testing Accuracy (%)**	
	EELM	**ELM**	**EELM**	**ELM**	**EELM**	**ELM**
1	1.123207	0.001248	60.42	62.50	75.86	51.03
2	1.107607	0.001248	95.83	77.67	**100.00**	53.24
3	1.232408	0.00312	100.00	85.83	**100.00**	52.55
4	1.216808	0.001248	97.92	89.50	**100.00**	58.48
5	1.294808	0.001872	97.92	91.25	**100.00**	60.55
6	1.435209	0.001872	100.00	92.08	**100.00**	63.17
7	1.372809	0.001248	100.00	94.67	**100.00**	63.59
8	1.450809	0.001872	100.00	95.75	**100.00**	70.48
9	1.60681	0.001872	100.00	96.92	**100.00**	68.69
10	1.638011	0.002496	100.00	97.92	**100.00**	76.97

ISBN: 978-1-61804-227-9

TABLE III MEAN TESTING EFFICIENCY AND PROCESSING TIME FOR ALGORITHMS – ELM, EELM WITH SIGMOID ACTIVATION FUNCTION

Hidden Neurons	Training Time(secs)		Training Accuracy (%)		Testing Accuracy (%)	
	EELM	ELM	EELM	ELM	EELM	ELM
1	1.232	0.0012	43.75	70.83	65.52	51.72
2	1.123	0.0006	95.83	77.75	**100.00**	47.31
3	1.170	0.0006	97.92	85.58	**100.00**	52.41
4	1.342	0.0006	97.92	91.50	**100.00**	60.97
5	1.498	0.0006	100.00	92.75	**100.00**	61.24
6	1.435	0.0006	100.00	95.17	**100.00**	66.76
7	1.513	0.0031	100.00	97.50	**100.00**	75.86
8	1.576	0.0031	100.00	97.00	**100.00**	80.00
9	1.747	0.0019	100.00	99.75	**100.00**	85.10
10	1.950	0.0019	100.00	99.92	**100.00**	91.03

The generalization performance of classifier is evaluated in terms of classification accuracy on the training and validation data set. The training accuracy and testing accuracy for varying number of neurons and activation function chosen as radial basis for the ELM and EELM algorithm is plotted Fig. 3and Fig.4. Similarly training and testing accuracy for sigmoid activation function for the ELM and EELM algorithm is plotted in Fig. 5 and Fig. 6. From these figures, it is observed that optimizing the input weights EELM algorithm has a better generalization performance with a compact network. The mean training and testing efficiency, testing time for the algorithms ELM , EELM with two chosen different activation functions namely the radbas and sigmoid are tabulated in table II and table III. The observation reveals that the EELM algorithm takes a longer processing time than the ELM due to the iterations performed in the optimization of the network weights.

IV. CONCLUSION:

In this work, an attempt had been made to classify the pulmonary obstructive disease between the normal and abnormal subjects using Single Layer Feed forward Neural networks (SLFNs). The performance of ELM, a fast learning algorithm achieved a generalization of 91.03% of accuracy in 0.0019secs . The output weights in this algorithm are computed based on the prefixed input weights and hidden biases using the Moore-Penrose (MP) generalized inverse. Thus due to the random selection of number of hidden neurons, there may exist a set of non-optimal input weights and hidden biases. A genetic algorithm based on global search optimization called Evolutionary Extreme Learning Machine is observed to achieve a better generalization with a reduced network size. This classifier can also be extended to incomplete data prediction and classification thereby supporting the clinicians in the near accurate diagnosis and treatment of pulmonary disorders.

REFERENCES

[1] Jabłoński, J. Mroczka,"The problem of measurement data complexity for example of the general model of the central respiratory generator and recurrent plots analysis," Metrol. Meas. Syst., vol. XV, no. 4, 2008, pp. 457 - 472.

[2] A. G. Polak, D. Wysoczański, J. Mroczka "A multi-method approach to measurement of respiratory system mechanics," Metrol. Meas. Syst., vol. XIII, no. 1, 2006, pp. 3-17.

[3] Global Status report on Noncommunicable Disease, Available: http://www.who.int/nmh/publications/ncd_report_full_en.pdf

[4] Mahesh Veezhinathan & Swaminathan Ramakrishnan,"Detection of Obstructive Respiratory Abnormality Using Flow–Volume Spirometry and Radial Basis Function Neural Networks, J Med Syst, 2007, 31:461–465

[5] Luigi G. Franciosia et.al. "Markers of disease severity in chronic obstructive pulmonary disease," Pulmonary Pharmacology & Therapeutics 19 ,2006, 189–199.

[6] American Thoracic Society/European Respiratory Society. Definition, Diagnosis and Staging of COPD. Available: http://www.thoracic.org/COPD/1/definitions.asp

[7] T. L. Croxton, G. G. Weinmann, R. M. Senior, R.A. Wise, J. D. Crapo, A. S. Buist, " Clinical research in chronic obstructive pulmonary disease: needs and opportunities," Am J Respir Crit Care Med, 2003, 167(8), 1142–9.

[8] P. M. Calverley, P. Walker, " Chronic obstructive pulmonary disease," Lancet, 2003, 362(9389), 1053–61.

[9] Atsuyoshi Kawagoshi,,Noritaka Kiyokawa,Keiyu Sugawara,Hitomi Takahashi, Shunichi Sakata, Saori

Miura, Sachie Sawamura, Masahiro Satake, Takanobu Shioya, "Quantitative assessment of walking time and postural change in patients with COPD using a new triaxial accelerometer system," International journal of COPD, August,2013.

[10] David Price,Mark Small, Gary Milligan, Victoria Higgins, Esther Garcia Gil, Jordi Estruch, " Impact of night-time symptoms in COPD: a real-world study in five European countries, International journal of COPD, November,2013

[11] Valipour A, Lavie P, Lothaller H, Mikulic I, Burghuber OC. Sleep profile and symptoms of sleep

disorders in patients with stable mild to moderate chronic obstructive pulmonary disease. Sleep Med. 2011;12(4): 367–372.

[12] Budhiraja R, Parthasarathy S, Budhiraja P, Habib MP, Wendel C, Quan SF. Insomnia in patients with COPD. Sleep. 2012;35(3):369–375.

[13] Javier de Miguel Díez, Jorge Chancafe Morgan, Rodrigo Jiménez García, "The association between COPD and heart failure risk: a review," International Journal of COPD, June, 2013

[14] Katja Oshaug, Peder AHalvorsen, Hasse Melbye, " Should chest examination be reinstated in the early diagnosis of chronic obstructive pulmonary disease?", International Journal of COPD, July, 2013.

[15] Bapoje SR, Whitaker JF, Schulz T, Chu ES, Albert RK, "Preoperative Evaluation of the Patient with Pulmonary Disease," Chest 2007;132:1637-1645.

[16] Vinaya S Karkhanis, JM Joshi, "Spirometry in Chronic Obstructive Lung Disease(COPD), " Supplement to JAPI, February 2012 , vol. 60 23

[17] Thomas A Barnes,Len Fromer, "Spirometry use: detection of chronic obstructive pulmonary disease in the primary care setting," Clinical interventions in Aging, January, 2011.

[18] Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global strategy for the diagnosis, management and prevention of chronic obstructive pulmonary disease, 2009. Available : http://www.goldcopd.org.

[19] Anthonisen NR, Wright EC, Hodgkin JE, " Prognosis in chronic obstructive pulmonary disease", Am Rev Respir Dis. 1986;133(1):14–20.

[20] Burrows B, "The course and prognosis of different types of chronic airflow limitation in a general population sample from Arizona: comparison with the Chicago "COPD" series," Am Rev Respir Dis. 1989;140 (3 Pt 2):S92–S94.

[21] Walters JA, Walters EH, Nelson M, et al, "Factors associated with misdiagnosis of COPD in primary care", Prim Care Respir J. 2011, 0(4): 396–402.

[22] Averame G, Bonavia M, Ferri P, et al; "Alliance Project" Study Group. Office spirometry can improve the diagnosis of obstructive airway disease in primary care setting. Respir Med. 2009;103(6):866–872.

[23] Christian Ghattas, Allen Dai, David J Gemmel, Magdi HAwad, " Over diagnosis of chronic obstructive pulmonary disease in an underserved patient population," International Journal of COPD, November, 2013.

[24] Wai Cho Yu, Sau Nga Fu, Emily Lai-bun Tai, Yiu Cheong Yeung, Kwok Chu Kwong, Yui , Chang , Cheuk Ming Tam, Yuk Kwan Yiu, " Spirometry is underused in the diagnosis and monitoring of patients with chronic obstructive pulmonary disease (COPD)" International Journal of COPD, August, 2013.

[25] Guillermo M Albaiceta, Esteban Garcia and Franciso taboada, " Comparative study of four sigmoid model of pressure- volume curve in acute lung injury," Biomedical Engineering online, 2007.

[26] Guang-Bin Huang, Qin-Yu Zhu, Chee-Kheong Siew, "Extreme learning machine: Theory and applications", Neurocomputing, 70 2006, 489–501.

[27] Dianhui Wang, Guang – Bin Huang , "Protein sequence classification using extremem learning machine," Proceeding of international conference on Neural networks, Canada, 2005.

[28] S. Suresh, S. Saraswathi, N. Sundararajan, "Performance enhancement of extreme learning machine for multicategory sparse data classification problems," Engineering applications of artificial intelligence, 2010.

[29] Luciano D. S. Pacifico and Teresa B. Ludermir, "Improved Evolutionary Extreme Learning Machines Based on Particle Swarm Optimization and Clustering Approaches", International Journal Of Natural Copmputing Research, 2012.

Zhu, Q.-Y.; Qin A.K., Suganthan P.N. & Huang G.-B. (2005), "Evolutionary Extreme Learning Machine, Pattern Recognition", Vol. 38, pp. 1759-1763.

[31] R.Storn, K.Price, "Differential Evolution – A simple and Efficient Heuristic for Global Optimization over Continuous Spaces", Journal of global optimization, vol.11, 1997, 341-359.

ISBN: 978-1-61804-227-9

Preliminary Design of Seismic Isolation Systems Using Artificial Neural Networks

Samer A. Barakat and Mohammad H. AlHamaydeh

Abstract— This works attempts to implement artificial neural networks (ANN) for modeling Seismic-Isolation (SI) systems consisting of Natural Rubber Bearings and Viscous Fluid Dampers subject to Near-Field (NF) earthquake ground motion. Fourteen NF earthquake records representing two seismic hazard levels are used. The commercial analysis program SAP2000 was used to perform the Time-History Analysis (THA) of the MDOF system (stick model representing a realistic five-story base-isolated building) subject to all 14 records. Different combinations of damping coefficients (c) and damping exponents (α) are investigated under the 14 earthquake records to develop the database of feasible combinations for the SI system. The total number of considered THA combinations is 350 and were used for training and testing the neural network. Mathematical models for the key response parameters are established via ANN. The input patterns used in the network included the damping coefficients (c), damping exponents (α), ground excitation (peak ground acceleration, PGA and Arias Intensity, I_a). The network was programmed to process this information and produce the key response parameters that represent the behavior of SI system such as the Total Maximum Displacement (D_{TM}), the Peak Damper Force (P_{DF}) and the Top Story Acceleration Ratio (TSAR) of the isolated structure compared to the fixed-base structure. The ANN models produced acceptable results with significantly less computation. The results of this study show that ANN models can be a powerful tool to be included in the design process of Seismic-Isolation (SI) systems, especially at the preliminary stages.

Keywords— seismic, base-isolation, supplemental viscous damping, near-field, ANN.

I. Introduction

The main objective of designing BI system for any structure is to guarantee the functions for which it was built by maintaining its functionality and its structural integrity. The structural control is based on models that used to predict the behavior of a structure. In general, it is interesting to find out how changes in the input variables affect the values of the response variables. ANN approaches have potential value for predicting the behavior of the SI systems.

Seismic-Isolation (SI) with and without supplemental damping for energy dissipation has proven to be an effective method of control for structures during seismic events. Many researchers investigated the use of different types of isolators see for example [1]-[14]. When faced with the challenge of limiting the Total Maximum Displacement (DTM) to practical limits, especially in NF sites, often times the designer would rely on Viscous Fluid Dampers (VFDs). Once supplemental damping is deemed necessary, many designers would prefer utilizing the linear behavior of NRB isolators combined with the supplemental damping provided by VFDs. This system has also been used in many projects in the USA and Japan [15]-[20]. This work utilized ANN to model the behavior of the combined NRB-VFD system under dual-level ensembles of 14 NF earthquake motions. To develop the database of feasible combinations for the SI system, the overall SI system performance is evaluated, for different combinations of damping coefficients (c) and damping exponents (α), under the 14 earthquake records. For that purpose, a Multi-Degree-Of-Freedom (MDOF) system is adopted and the commercial analysis program SAP2000 [21] was used to perform the Time-History Analysis (THA). The key response parameters considered are the DTM, the Peak Damper axial Force (PDF) and the Top Story Acceleration Ratio (TSAR) of the isolated structure compared to the fixed-base structure. The total number of considered THA combinations is 350 and were used for training and testing the neural network. The input patterns used in the network included the damping coefficients (c), damping exponents (α), ground excitation (peak ground acceleration, PGA and Arias Intensity, I_a). Mathematical models for the key response parameters are established via ANN.

II. Model description

A. The damping system

The fluid viscous damper force-velocity behavior is governed by the mathematical expression described in Eqn (1):

$$F_D = c\,\mathrm{sgn}(\dot{v})\left|\dot{v}\right|^{\alpha} \tag{1}$$

where FD is the damper force (kN), c is the damping coefficient (kN-(s/m)$^{\alpha}$), v is the damper extensional velocity (m/s), α is the velocity exponent (for a linear damper, α =1) and sgn denotes signum function describing the velocity sign. In seismic applications, nonlinear dampers with damping exponent less than unity are preferred due to their softening or

Samer A. Barakat is a professor at the Department of Civil and Environmental Engineering at the University of Sharjah, Sharjah, UAE (e-mail: sbarakat@sharjah.ac.ae).

Mohammad H. AlHamaydeh is an associate professor at the Department of Civil Engineering at the American University of Sharjah, Sharjah, UAE (phone: +971-6-515-2647; fax: +971-6-515-2979; e-mail: malhamaydeh@aus.edu).

ISBN: 978-1-61804-227-9

yielding nature at higher velocities and the stiffening effect at lower velocities. This nonlinear characteristic results in significant reduction of base displacement in response to strong ground shaking, particularly in NF situations. Furthermore, it puts practical limitations on the amount of force transferred to the structural elements. Although some manufacturers can produce dampers with a value as low as 0.1, the typically used values range from 0.4 to 0.7. In this study, the damping exponent values considered range from 0.4 to 1.0 with intermediate values equally spaced at 0.15 intervals. Except when α=1, the damper elements are behaving nonlinearly and they are the only part of the model that could exhibit nonlinearity. The superstructure of an isolated structure is typically expected to remain near-elastic throughout significant seismic events, which justifies the linear analysis here. The considered values for the damping coefficient c range from 175 to 525 kN-(s/m)$^{\alpha}$ with intermediate values equally spaced at intervals of 88 kN-(s/m)$^{\alpha}$. The five different c and α values are used to generate 25 combinations to be investigated in the analysis.

B. *SAP2000 modeling of the MDOF system*

To include higher modes influence in the base isolation behavior, MDOF systems should be considered. The numerical analysis of the proposed MDOF system was conducted using the commercial finite element code SAP2000 (CSI 2010) [20]. For this study, a simple lumped-mass stick model is used to represent a five-story base-isolated building which has been introduced by Kelly et al. [22]. The building structural parameters and isolator properties are proportioned such that the fundamental period of vibration is 2.5 s and the modal damping is 5% of critical. The MDOF system has been modeled in three different configurations (boundary

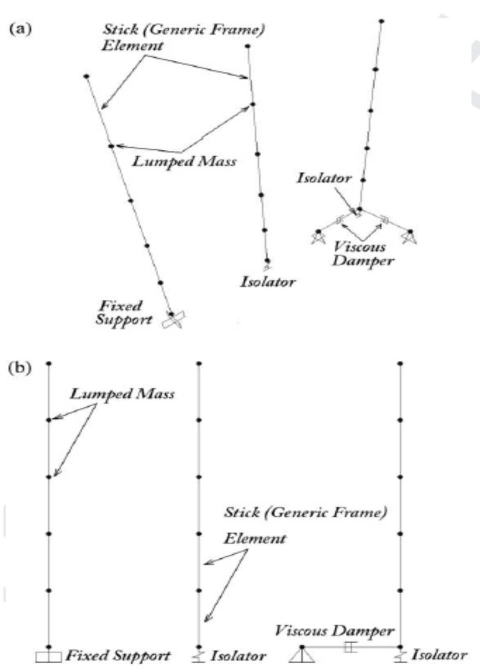


Fig. 1. SAP2000 MDOF Model

conditions) for comparison: (1) fixed base, (2) isolated without dampers, and (3) isolated with dampers. Figure 1 shows the SAP2000 model used to model the MDOF system.

C. *The ground excitation*

There are different parameters to characterize and quantify earthquake demand and damage potential. The Peak Ground Acceleration (PGA) and Arias Intensity (I_a), first introduced by Arias [23], are good examples of such parameters. The Arias Intensity (Ia) is adopted here as the main descriptor of the ground motion excitations for its ability to capture the earthquake amplitude variation, frequency content and duration. The Arias Intensity can be computed using the following Eqn (2):

$$I_a = \frac{\pi}{2g} \int_0^t a(t)^2 dt \qquad (2)$$

where $\alpha(t)$ is the ground acceleration history in g's, and g is the gravitational acceleration.

III. Results and discussion

A. *Time-history analysis*

The commercial analysis program SAP2000 was used to perform the Time-History Analysis (THA) of the MDOF system subject to all 14 records. The analyses were performed for all three MDOF systems representing the benchmark fixed-base, as well as the isolated buildings with and without dampers. For each of the records, three key response parameters were considered: the DTM, the Peak Damper Force (PDF) and the TSAR at the isolated and the fixed structures. The results for the three response parameters when the 25 different combinations of damping coefficients and damping exponents are investigated under the 14 earthquake records consist of 350 (5 c values×5 α values×14 records) THA combinations. The THA data set is further divided into two subsets: Set 1, consisting of THA results from 300 (c, α, EQ) combinations is used to produce mathematical models via MRA and ANN; Set 2, consisting of THA results from 50 (c, α, EQ) combinations is used to test the developed mathematical models. It should be emphasized here that the data from 50 testing combinations were not included in the modeling phase.

B. *Neural Network (NN) Analysis*

The primary objective of this section is to present a computer oriented method based on artificial neural networks (ANNs) technology to assess the structural behavior of BI systems. It is basically made up of a computer empirical model that maps the output variable or target value (DTM, PDF, TSAR) directly to a set of input variables (C, α, PGA, Ia) thus deducing a functional relationship for prediction purposes. The data utilized in this study were initially generated by a commercial analysis program SAP2000. The advantage of proposed ANNs model over *SA*P2000 computer programs is that it requires minimal

ISBN: 978-1-61804-227-9

input data with minor intermediate computations.

Neural networks analysis is an information processing technique in which a neuron is the main element and it is an operator with inputs and outputs, associated with a transfer function, f , called a "sigmoid" interconnected by synaptic connections or weights, w (plus a bias). Fig. 2 illustrates how information is processed through a single neuron. The way in which inputs are combined, how the resulting internal activation level is used to produce an output, and the way in which weights and biases are changed (the learning rule) are collectively called network architecture (see Fig. 3) results in a network paradigm. Several network paradigms are commonly involved with each designed to solve specific kinds of problems.

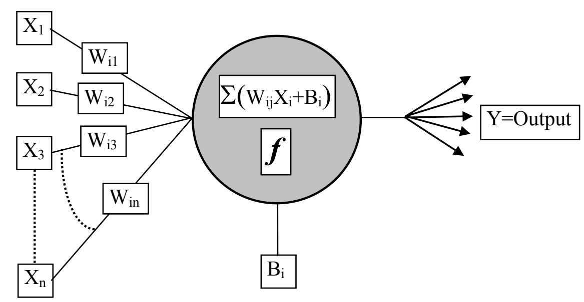


Fig. 2. Schematic of a single neuron

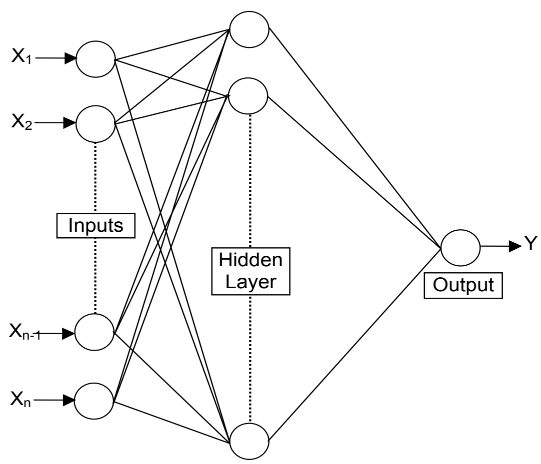


Fig. 3. Typical neural networks architecture

C. *Network Model and Architecture*

The mapping of the inputs to the outputs (target values) is established through the neural net architecture. For the problem considered here, the following architecture was considered: feed-forward network with 5 hidden neurons, 3 output neurons, tansig hidden neurons and linear output neurons. Weights and biases joining the input nodes to the hidden nodes, and those bridging this latter node to the output node were initially assigned randomly see Fig.4.

D. *Neural Networks Solution of BI Systems*

The NN model considered consists of an input layer with 4 input parameters to represent the (C, α, PGA, Ia)), an output layer with 3 output parameters to represent the (DTM, PDF, TSAR) and one hidden layer with 5 neurons, as can be seen in

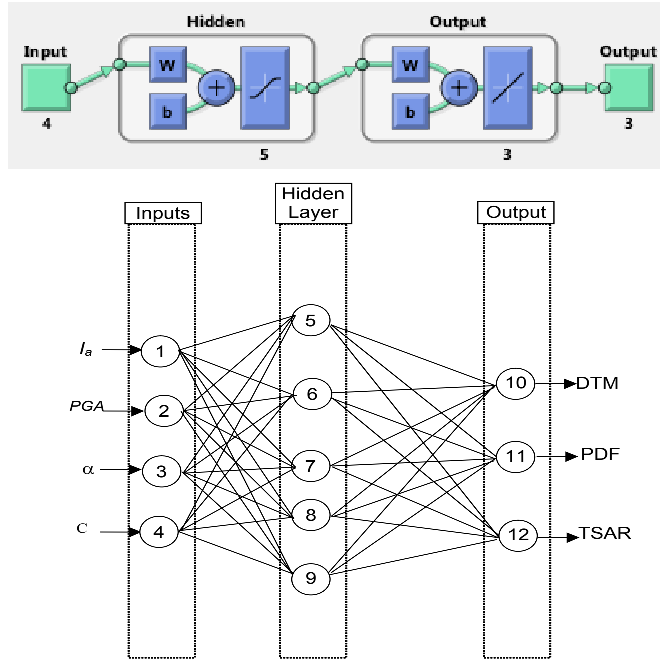


Fig. 4. Neural networks architecture

Fig. 4. Every neuron in the network is fully connected with each neuron of the next layer. The most appropriate model was sought by training the artificial neural networks with 3, 5, and 7 hidden nodes. The progress of the networks' training was monitored by observing the output error after each training cycle. The results showed that the average sum squared error decreased with increasing number of hidden nodes. Fig. 5 shows the training progress of the final network with 5 hidden nodes. The asymptotic shape of the curve implies that the network learning was notably complete by the end of the training. Furthermore, Fig. 5 indicates that approximately 51 Epochs were required for convergence.

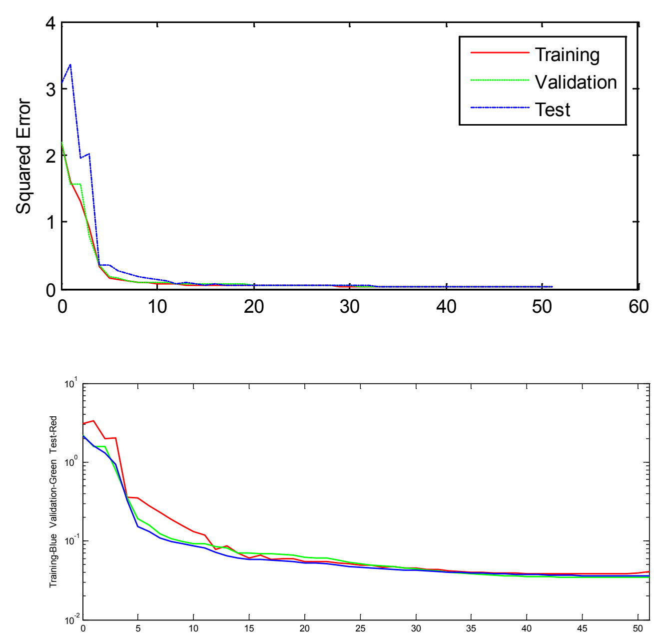


Fig. 5. Performance of the Neural network

E. Recall and Prediction Accuracy

The accuracy of the adopted NNs model was first checked by recalling the same data initially use to train it (set 1). Once the model was deemed acceptable, its prediction accuracy was tested against a new generated set of data (set 2). It should be stressed that all of the data in this latter set were initially withheld from the neural networks. In a similar fashion as in the recall test, the input values from these sets of data were presented to the model to perform the necessary calculations and produce corresponding outputs. Figs. 6 to 8 respectively show a comparison between the theoretical SAP2000 results and the recalled values by the ANN models.

For the models adopted in this work, the prediction accuracy is investigated. Data (Set 2) which consist of 50 randomly selected combinations is used to perform three ANN prediction tests. As mentioned earlier, all of the data in this testing set was initially withheld from the ANN. The results of these tests are shown in Figures 9 to 11. The closeness of the points to the equality line serves only to indicate the validity of the ANN models.

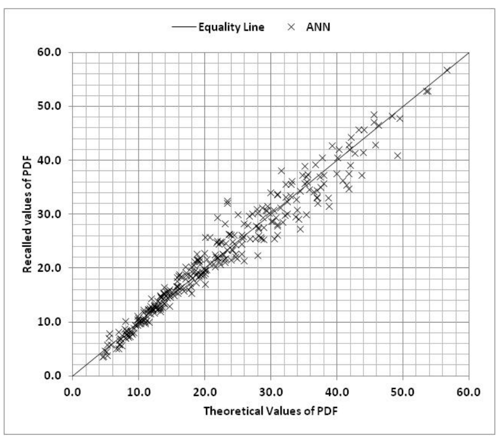


Fig. 6. Recalled PDF values by ANN vs. theoretical data (set1)

IV. SUMMARY AND CONCLUSION

The feasibility of using ANN to model and predict the dynamic behavior of Seismic-Isolated (SI) systems was investigated. THA was performed using SAP2000 for three MDOF systems representing a typical seismic isolated structure with a natural period of vibration equal to 2.5 s. Two ensembles of seven ground motion records representing two hazard levels (DBE and MCE) and return periods (475 and 950 years) were used.

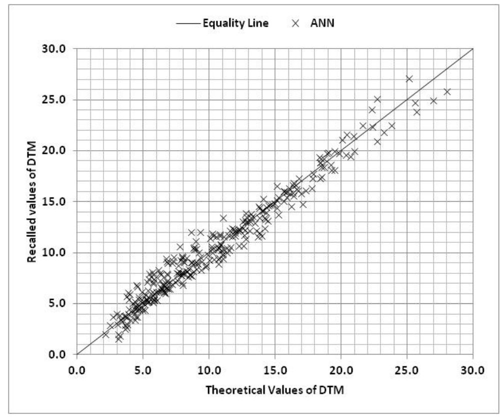


Fig. 7. Recalled DTM values by ANN vs. theoretical data (set1)

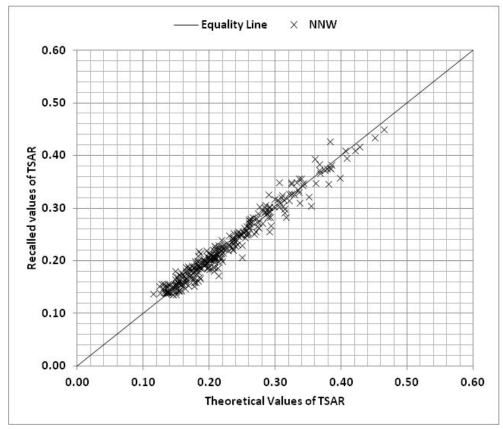


Fig. 8. Recalled TSAR values by ANN vs. theoretical data (set1)

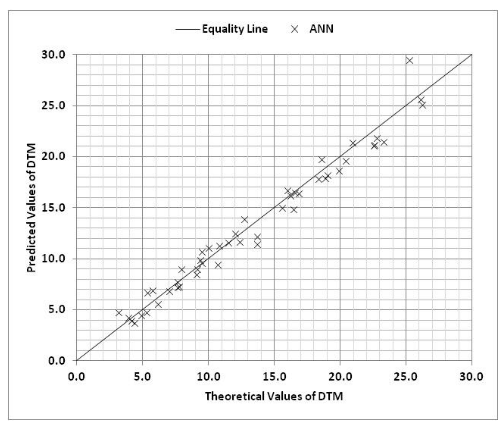


Fig. 9. Predicted PDF values by ANN vs. theoretical data (set2)

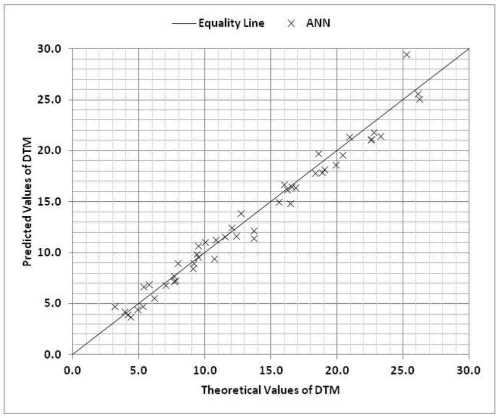


Fig. 10. Predicted DTM values by ANN vs. theoretical data (set2)

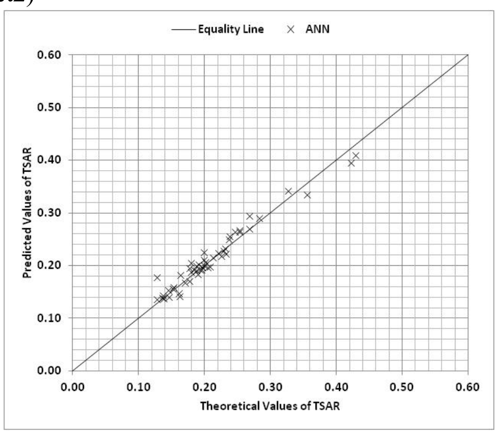


Fig. 11. Predicted TSAR values by ANN vs. theoretical data (set2)

The range of the SI system properties covered several feasible solutions comparable to the state-of-the-practice designs. Three key response parameters were selected to be modeled using MRA, namely; DTM, PDF, and TSAR of the isolated structure compared to the fixed-base structure. The response parameters, as well as the characteristics of the ground motions were utilized to develop several ANN models. For each of the key response parameters, the best fitting ANN models were selected. The design process of SI systems is iterative, complex and requires considering many feasible alternatives. Moreover, the most widely used analysis tool, the nonlinear THA, is very expensive in terms of CPU time which adds another layer of complexity to the situation. Therefore, simplifying techniques are extremely valuable especially at the preliminary design stages. In this investigation, it was demonstrated that the ANN modeling is a strong candidate to accompany, if not replace, the nonlinear THA. Once ANN models are developed through performing rigorous nonlinear THA, such as the presented work, several design options of SI systems can be easily selected and compared. Since the use of ANN models to evaluate the key response parameters is significantly simpler than performing THA, much more feasible solutions can be readily investigated and compared. This can be particularly valuable in the early design stages of SI systems utilizing the NRB-VFD combination.

REFERENCES

[1] M. AlHamaydeh, S. Barakat, F. Abed, "Multiple Regression Modeling of Natural Rubber Seismic-Isolation Systems with Supplemental Viscous Damping for Near-Field Ground Motion," Journal of Civil Engineering and Management, 19 (5), 665–682, 2013.

[2] R. Skinner, W. Robinson, G. McVerry, "An introduction to seismic isolation", London: John Wiley and Sons. GH, 1993.

[3] J. Kelly, "Earthquake-Resistant Design with Rubber", Springer-Verlag, Second Edition, London Limited, 1997.

[4] F. Naeim and J. Kelly, "Design of Seismic Isolated Structures", John Wiley & Sons Inc., 1999.

[5] R. Jangid, "Optimum friction pendulum system for near-fault motions," *Engineering Structures*, 27(3): 349–359, 2005.

[6] V. Matsagar, R. Jangid, "Viscoelastic damper connected to adjacent structures involving seismic isolation," *Journal of Civil Engineering and Management*, 11:4, 309-322, 2005.

[7] A. Rodriguez-Marek, "Near fault seismic site response", Ph.D. thesis. Berkeley: Civil Engineering, University of California, p. 451, 2000.

[8] R. Jangid and J. Kelly, "Base isolation for near-fault motions," *Earthquake Engineering & Structural Dynamics*, 30(5): 691-707, 2001.

[9] A. Chopra, C. Chintanapakdee, "Comparing response of SDF systems to near fault and far-fault earthquake motions in the context of spectral regions*," Earthquake Engineering and Structural Dynamics*, 30: 1769–89, 2001.

[10] G. Macrae, D. Morrow, C. Roeder, "Near-fault ground motion effects on simple structures," *ASCE Journal of Structural Engineering*, 127(9):996–1004, 2001.

[11] N. Makris, C. Black, "Dimensional analysis of bilinear oscillators under pulse-type excitations," *ASCE Journal of Engineering Mechanics*, 130(9):1019–31, 2004.

[12] C. Providakis, "Effect of LRB isolators and supplemental viscous dampers on seismic isolated buildings under near-fault excitations," *Engineering Structures* 30(5):1187–1198, 2008.

[13] C. Providakis, "Effect of supplemental damping on LRB and FPS seismic isolators under near-fault ground motions," *Soil Dynamics and Earthquake Engineering*, 29(1): 80–90, 2009.

[14] A. El-Sinawi, M. AlHamaydeh, A. Jhemi, "Optimal Control of Magnetorheological Fluid Dampers for Seismic Isolation of Structures," Mathematical Problems in Engineering, Vol. 2013, Article ID 251935, 2013. DOI:10.1155/2013/251935.

[15] S. Hussain, "Performance of Base Isolated Buildings in the Northridge Earthquake, Seismic Base Isolation - State of the Practice Seminar," Structural Engineers Association of Southern California, 1994.

[16] S. Hussain, E. Retamal, "A Hybrid Seismic Isolation System - Isolators with Supplemental Viscous Dampers," Proceedings of the First World Conference on Structural Control; International Association for Structural Control, Los Angeles, 3-5, volume 3, pages FA2-53--FA2-62, August 1994.

[17] S. Hussain, M. Al Satari, "Design of a Seismic Isolation System with Supplemental Viscous Damping for a Near-Fault Essential Services Facility," Proceedings of the 76th SEAOC Annual Convention, Squaw Creek, 2007.

[18] S. Hussain, M. Al Satari, "Innovative Design of a Seismic Isolation Supplemental Viscous Damping Systems of an Essential Services Facility in a Near-Fault Region," Proceedings of the 14th World Conference on Earthquake Engineering, Beijing, China, 2008.

[19] M. Al Satari, J. Abdalla, "Optimization of A Base-Isolation System Consisting of Natural Rubber Bearings and Fluid Viscous Dampers," Proceedings of the 11th World Conference on Seismic Isolation, Energy Dissipation and Active Vibration Control of Structures, Guangzhou, China, 2009.

[20] M. AlHamaydeh, S. Hussain, "Innovative Design of a Seismically-Isolated Building with Supplemental Damping," Proceedings of 14th European Conference on Earthquake Engineering (14ECEE), Ohrid, Republic of Macedonia, 2010.

[21] Computers and Structures, Inc., SAP2000 Reference Manual. Berkeley, California, 2010.

[22] J. Kelly, G. Leitmann, A. Soldatos, "Robust control of base-isolated structures under earthquake excitation*," Journal of Optimization Theory & Applications*, Vol.53, No.2, May 1987.

[23] Arias, A., "A Measure of Earthquake Intensity," Seismic Design for Nuclear Power Plants, MIT press, Cambridge, MA. pp 438-483, 1970.

The Use of the Modified Semi-bounded Plug-in Algorithm to Compare Neural and Bayesian Classifiers Stability

Ibtissem Ben Othman and Faouzi Ghorbel

Abstract**—**Despite of the widespread use of the neural networks in the industrial applications, their mathematical formulation remains difficult to analyze. This explains a limited amount of work that formally models their classification volatility. Referring to the statistical point of view, we attempt in this work to evaluate the classical and Bayesian neural networks stability degree compared to the statistical methods stressing their error rate probability densities. The comparison based on this new criterion is performed using the modified semi-bounded Plug-in algorithm.

Keywords**—**Bayesian neural networks, error rate density, modified semi-bounded Plug-in algorithm, stability.

I. Introduction

As the samples have a limited size, the classification in high dimension spaces remains one of the essential problems for pattern recognition. Therefore, dimension reduction is often required in the first step. The Artificial neural networks (ANNs or NNs) have been commonly used to classify the data and to solve any non-linear dimension reduction. Actually, the ANNs partially substituted the statistical methods in the industrial field. The Traditional statistical classification methods are based on the Bayesian decision rule, which presents the ideal classification technique in terms of the minimum of the probability error. However, in the non parametric context, applying Bayes classifier requires the estimation of the conditional probability density functions. It is well known that such task needs a large samples size in high dimension. The Principal Components Analysis (PCA) and the Linear Discriminant Analysis (LDA) are generally used to reduce the dimension of the feature space. They are applied to the original feature space in order to select a limited number of discrimination directions before applying non parametric Bayes classifier [26]. While PCA seeks for efficient representation directions, the Fisher LDA tries to find efficient discrimination directions. The LDA is commonly preferred over the PCA. In fact, the LDA is able to recognize the different classes, whereas the PCA deals with the data without paying any particular attention to the underlying class [1], [26].

Many authors carried out many comparison studies of neural and statistical classifiers. A recent review of these studies is presented in [14]. It aims to give a useful insight into the neural and statistical methods capabilities. These methods have widely been used to solve complex problems and have proven quite successful in many applications, as the dimension reduction and classification problems. Tam and Kiang compare in [12] the neural networks and the linear classifiers (Discriminant Analysis, logistic regression and k Nearest Neighbour) for bank bankruptcy prediction in Texas. They showed that ANNs offered better predictive accuracy than other classifiers. Patuwo and al evaluate, in [27], the neural networks performance against discriminant analysis for some classification problems. They proved that neural approaches are comparable but not better than the LDA in two-group two-variable problems.

In order to compare the neural and statistical techniques, most of researchers try to compare their accuracy prediction while forgetting the instability criterion of NNs . This paper studies the stability of different network classifier results compared to the statistical methods. By estimating the error rate probability density function (pdf) of each classifier, we evaluate their stabilities. In order to estimate its pdf, we apply the Plug-in kernel algorithm, which optimizes its smoothing parameter. The miss classification error is positive value, so we opt for the modified semi-bounded Plug-in algorithm to improve the pdf estimation precision since pdf support information is known.

So, the present work will be organized as following: First, we start by presenting the main topic and briefly introduce the neural approaches. Here we deal with the Bayesian approach for the artificial neural networks. Then we lead a comparative study between the neural and the statistical approaches. Here we focus on the stability degree and visualize the results through stochastic simulation of particular distributions (for example: Gaussian, Gamma, Beta, etc…). Finally, we intend to test the classifiers stability and performance for the handwritten digits recognition problem.

Ibtissem Ben Othman is with the School of Computer Sciences, CRISTAL Laboratory, GRIFT Research Group, Manouba 2010, Tunisia. (e-mail: ben.othman.ibtissem@ gmail.com).

Faouzi Ghorbel is a professor at the School of Computer Sciences and the director of CRISTAL Laboratory, GRIFT Research Group, Manouba 2010, Tunisia. (e-mail: faouzi.ghorbel@ensi.rnu.tn).

ISBN: 978-1-61804-227-9

II. Neural Approaches

In pattern recognition, the neuronal networks can be categorized into three different types, depending on their application objectives (size reduction, classification or both). The first networks category is the features extractors NNs. It aims to reduce the learning set dimensionality thereby extracting the relevant primitives. The classifier NNs is the second networks category. Its main duty is to classify the extracted features regardless of the dimensionality. The third networks category is the mixed NNs, which present a combination of both defined types. Once the first networks layers carry out the primitive extraction, the last layers classify the extracted features. An interesting example is the multilayer NNs that uses the back-propagation algorithm. Thanks to its several hidden layers, the multilayer perceptron (MLP) can reduce non-linearly the data dimension and extract its relevant characteristics. Finally, a linear separation is applied to classify these extracted primitives in the output layer. Based on the results from [19] and [11], a MLP with one hidden layer is generally sufficient for most problems including classification. Therefore, all used networks, in our study, will have a unique hidden layer. The number of neurons in this layer could only be determined by experience and no rule is specified.

However, the number of nodes in the input and output layers is set to match the number of input and target parameters of the given process, respectively. Thus, the NNs have a complex architecture and designing the optimal model for such application is not so easy.

By estimating the weights matrices, the training algorithm aims to reduce the difference between the ANN outputs and the known target values, such that an overall error measure is minimized. The most commonly used performance measure is the mean squared error (MSE) defined as:

$$MSE = \frac{1}{N}\sum_{j=1}^{N}\left(t_j - y_j\right)^2 \qquad (1)$$

where t_j and y_j represent the target and network output values for the j^{th} training sample respectively, and N is the training samples size. The NNs training algorithms, for the classification, are mostly employed in a supervised learning process. In the proposed technique, improvements will be required for MLP with the back-propagation algorithm.

A. Neural Networks limitations

The ANN produces a black box model in terms of only crisp outputs, and hence cannot be mathematically interpreted as in statistical approaches. The most common representation mode of the output layer in pattern recognition is the "local representation" one. In this representation, each output neuron represents one of the classes to which samples can belong [14]. The MLP desired outputs provide a value close to 1 for the class neuron of the considered object and values close to zero for the other nodes. However, the ANNs outputs values are not similar to those desired, and they can even be negative values. For a discrete representation {0,1} of output neurons, Jolliffe normalizes, in [9], the obtained outputs, and considers the new normalized values as a posterior probability. In [7], the authors have used the Softmax transfer function, which ensures that the ANN outputs are homogeneous to a posterior probability. Till today, the quality of this approximation has never been proved. However, users of these networks are based on this approximation as a thresholding function to binarize the obtained outputs. This non proved approximation, the black box nature, the lack of control over its mathematical formulation and the non fixed architecture of the optimal NN model explain the instability of its classification results as against the statistical ones. Thus, after the training phase, small changes in the test samples can introduce a large variance in its prediction results. During the training phase, the NN classifier might learn the data very well in order to reach best results. As a consequence, this can lead to the NN instability; the overfitting may create a high variance while testing the new data. Indeed, the overfitting related problems got the attention of the literature and researches kept looking for suitable solutions. The classical one is the cross validation method [18], [21]. Combining several neural classifiers is another solution which may improve the performance and stability classification [8], [5], [13], [18]. The *bias plus variance* decomposition of the prediction error, introduced by German and al in [20], presents a solution for the overfitting problem. In order to reduce the overfitting effect of NNs, Mackay has proposed, in (Mackay, 1992), a probabilistic interpretation of neural networks learning methods, thereby using Bayesian techniques.

B. Bayesian Neural Networks

The Bayesian approach for NNs was originally developed by Mackay in [3] and reviewed by Bishop in [2] and Mackay in [4]. This approach has been devoted to improving the conventional NN learning methods while adding a penalty term to the classical error function. The resulting function is defined by:

$$S(w) = E_D(w) + \mu E_w(w) \qquad (2)$$

The penalty term $E_w(w) = \frac{1}{2}\sum_{i=1}^{m} w_i^2$ (where m is the total number of parameters) controls the model complexity. The main idea is to find the optimal value of the regularization coefficient μ that gives the best tradeoff between the overfitting and the underfitting problems. This optimal value can be found by using probabilistic interpretation of NN learning which controls automatically its complexity.

The Bayesian approach assigns a probability density function (pdf) to each NN parameter w_i (weights, biases, number of neurons, NN outputs, etc). This pdf is initially affected to a prior distribution, and once the data have been observed, it will be converted to a posterior distribution using Bayes theorem:

$$p(w|t,x) = \frac{p(t,x|w)\,p(w)}{p(t,x)} \qquad (3)$$

III. Approaches Comparison

Most of the current research almost exclusively uses the criterion of prediction accuracy, while evaluating classifiers and comparing their performance, but ignoring the fact that the NNs present unstable results. In this paper, a new approach to comparing neural and statistical classifiers is proposed. In addition to the prediction accuracy criterion, a stability comparison based on estimating classifiers error rate probability density functions is presented.

A. Error rate density estimation

We start first by training the two classifiers to be compared, and then measure the error rate produced by each classifier on each one of N independent test sets. Let $(X_i)_{1\le i\le N}$ be the N generated error rates of a given classifier (Bayes or ANN). These error rates $(X_i)_{1\le i\le N}$ are random variables having the same probability density function (pdf), $f_X(x)$. The $(X_i)_{1\le i\le N}$ are assumed to be independent and identically distributed.

We suggest to estimate the pdf of the error rates for each classifier using the kernel method proposed in [10] and [25], where the involved smoothing parameters h_N are estimated by optimizing an approximation of the integrated mean square error (IMSE). The kernel estimator of the probability density is defined as follows:

$$\hat{f}_N(x) = \frac{1}{Nh_N}\sum_{i=1}^{N} K\left(\frac{x - X_i}{h_N}\right) \tag{4}$$

In our study, $K(.)$ is chosen as the Gaussian kernel:

$$K(x) = \frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}x^2) \tag{5}$$

The choice of the optimal smoothing parameter h_N^* is very important. Moreover, Researchers have introduced different methods that minimize the integrated mean square error ($IMSE \approx \frac{M(K)}{Nh_N} + \frac{J(f(X))h_N^4}{4}$) to define the optimal bandwidth. The smoothing parameter h_N^* becomes as follows:

$$h_N^* = N^{-1/5}[J(f_X)]^{-1/5}[M(K)]^{+1/5} \tag{6}$$

where;

$$M(K) = \int_{-\infty}^{+\infty} K^2(x)dx$$

and

$$J(f_X) = \int_{-\infty}^{+\infty} (f_X''(x))^2 dx$$

The goodness of estimation depends on choosing an optimal value for the smoothing parameter. Calculating its optimal value, with a direct resolution of the equation (4), seems very difficult. We opt for the recursive resolution: The *Plug-in* algorithm. Actually, a fast variant of known conventional Plug-in algorithm has been developed [24]. It applies directly a double derivation of the kernel estimator analytical expression in order to approximate the function $J(f)$.

B. Modified semi-bounded Plug-in algorithm

The set of observed error rates $(X_i)_{1\le i\le N}$ of each classifier is a set of positive values. In this case, the kernel density estimation method is not that attractive. When estimating the probability densities, which are defined in a bounded or semi-bounded space $U \subset \Re^d$, we will encounter convergence problems at the edges : the Gibbs phenomenon. Several authors have tried to solve this issue and presented some methods to estimate the probability densities under topological constraints on the support. The orthogonal functions method and the kernel diffeomorphism method are two interesting solutions [22], [23]. The kernel diffeomorphism method is based on a suitable variable change by a C1-diffeomorphism. Although, it is important to maximize the value of the smoothing parameter in order to ensure a good estimation quality. The optimization of the smoothing parameter is performed by the Plug-in diffeomorphism algorithm which is a generalization of the conventional Plug-in algorithm [16].

For complexity and convergence reasons, we propose in this paper a modified semi-bounded Plug-in algorithm. This algorithm version is based on the variable change of the positive error rates: $Y = Log(X)$. In order to define new classification quality measure, a sequence of three steps is performed:

Step 1: using the variable change $Y = Log(X)$, the kernel estimator expression becomes:

$$\hat{f}_Y(y) = \frac{1}{Nh_N^*}\sum_{i=1}^{N} K\left(\frac{y - Y_i}{h_N^*}\right) \tag{7}$$

Step 2: iterate the conventional Plug-in algorithm for the transformed data.

Step 3: compute $\hat{f}_X(x) = \frac{\hat{f}_Y(Logx)}{x}$

The modified semi-bounded Plug-in algorithm produces a sufficient precision for the densities estimation and the stability aspects. It tends to be a good criterion for the stability comparison of the different classifiers and reduction algorithms.

C. Performance and stability comparison

Some classifiers are instable, small changes in their training sets or in constructions may cause large changes in their classification results. Therefore, an instable model may be too dependent on the specific data and has a large variance. In order to analyze and compare the stability and performance of each classifier, we have to illustrate their error rate probability densities in the same figure. While the probability density curve on the left has the small mean, the one on the right has the high mean. Clearly, the classifier, whose curve is on the left, is the most efficient one. An instable classifier is characterized by a high variance. When the variance is large, the curve is short and wide, and when the variance is small, the curve is tall and narrow. As a result, a classifier with the largest density curve is the least stable one. Therefore, a good

ISBN: 978-1-61804-227-9

model should find a balanced equilibrium between the error rate bias and variance. This criterion is basic for any stability and performance analysis of each classifier.

IV. SIMULATIONS

For the simulation phase, we propose a binary classification adapted to a mixture of two different distributions. We generate one train set including 1000 samples for each class. With this train set, we look to find the optimum transformation that represents the dimension reduction for both PCA and LDA methods before applying the Bayesian rule, and then to fix the optimal NN model parameters for both MLP and Bayesian NN.

In order to analyze the stability of the different approaches, we generate 100 supervised and independent test sets including 1000 samples for each class. For each test set, the classifier performance is evaluated by its error rate calculated from the confusion matrix. The error rate probability densities, retained for both approaches, are estimated using the modified semi-bounded Plug-in algorithm.

The comparison between the statistical and neural classifiers used in the present work (PCA-Bayes, Fisher-Bayes, MLP and Bayesian NN) is first summarized by the Gaussian mixture classification problem. Figure 1 shows the estimated error rate probability densities generated for the different classifiers on a mixture of two homoscedastic Gaussians (Figure1.a), two heteroscedastic Gaussians (Figure1.b), two superposed Gaussians (Figure1.c) and two truncated ones (Figure1.d). The stability and performance of the classifiers are also analyzed by presenting their error rate means and variances in table 1.

The first two cases ((a) and (b)) in figure 1 and table 1, show that the statistical classifiers (ACP-Bayes and Fisher-Bayes) are more efficient than the neural ones (they admit the smallest error rate means). However for the last complex cases of the two heteroscedastic superposed Gaussians (c) and the truncated ones (d), the error rate probability density functions of the neural models are on the left. We conclude that these models are the most efficient. Although, the neural approach remains the least stable classifier that presents the greatest variance and thus the widest curve. For these two complex cases, the linear reduction dimension methods (PCA and Fisher) fail to find the optimal projection subspace. Whereas, the Bayesian and classical NN perform well due to their non linear reduction dimension capability.

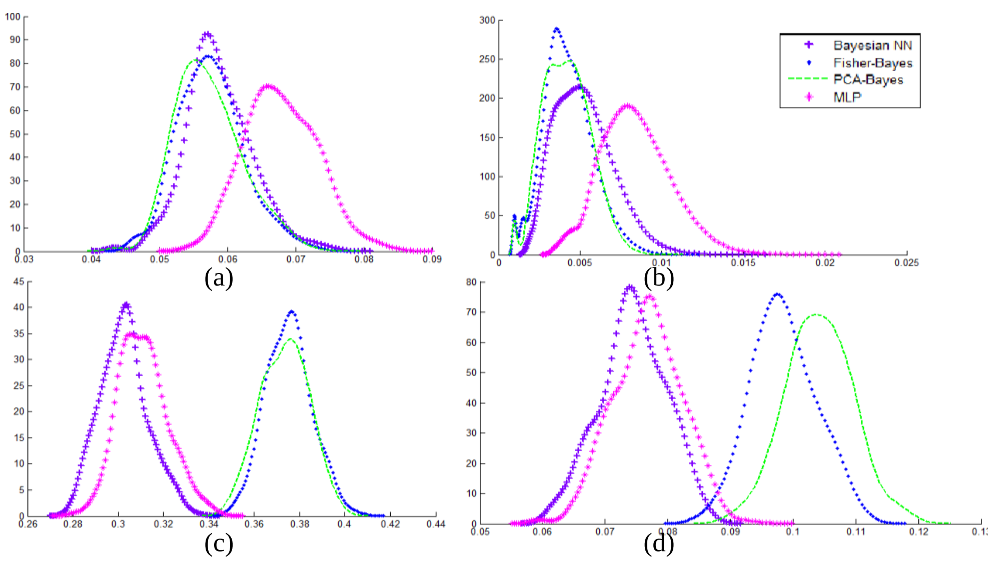


Figure 1: Error rate densities of PCA-Bayes (in green(--)), Fisher-Bayes (in blue(..)), MLP (in pink(*)) and Bayesian NN (in purple(+)).

Table 1: Comparison results of ANN and statistical classifiers.

Cases	Distributions		PCA-Bayes		Fisher-Bayes		MLP		Bayesian NN	
	Gaussian 1	*Gaussian 2*	*Mean*	*Variance*	*Mean*	*Variance*	*Mean*	*Variance*	*Mean*	*Variance*
a	$\mu_1=(1,..,1),\sum_1=$Identity	$\mu_2=(2,..,2),\sum_2=$Identity	0.0572	0.2287	0.0575	0.2223	0.0679	0.2697	0.0587	0.2284
b	$\mu_1=(0,..,0),\sum_1=$Identity	$\mu_2=(2,..,2),\sum_2=2*$Identity	0.0042	0.1721	0.0043	0.1972	0.0084	0.4231	0.0052	0.2590
c	$\mu_1=(0,..,0),\sum_1=$Identity	$\mu_2=(0,..,0),\sum_2=2*$Identity	0.3734	0.1054	0.3753	0.0983	0.3110	0.1140	0.3026	0.1094
d	$\mu_1=(0,0,0)$ $\sum_1=[0.06\ 0\ 0$ $0\ 0.01\ 0$ $0\ 0\ 0.01\ ]$	$\mu_2=(0.1,0.1,0.1)$ $\sum_2=[0.01\ 0\ 0$ $0\ 0.06\ 0$ $0\ 0\ 0.05\ ]$	0.1041	0.2804	0.0985	0.2612	0.0768	0.2877	0.0745	0.2709

To analyze the classifiers stability, the comparison study is often illustrated by a mixture of univariate distributions according to the Pearson System. This system presents a set of eight families of distributions, including Gaussian, Gamma and Beta ones. The Pearson system distributions are generally qualified by their four parameters; mean (μ), variance (σ^2), skewness (β_1) and kurtosis (β_2). Further details about the Pearson system can be found in [5] and [17].

The classifiers performances and stabilities are compared in the sense of their error rate means and variances in table 2. Figure 2 shows the estimated error rate densities generated for the different classifiers (the figure cases (a,b,…,f) correspond to the cases in table 2).

By analyzing the figures and the table above, we note that the neural approach admit the greatest variances, we can confirm that it is less stable than the statistical one.

Using the simulations concerning the different kinds of stochastic distributions, we illustrate the better stability of the Bayesian classifier against the neural one. In addition, the statistical approaches are proved to perform better than the neural networks in the simple classification problems. However, the results prove that the neural classifier performs better if the classification task tends to become complex (the truncated and the superposed distributions for the Gaussian simulations and the Pearson System distributions). Although the Bayesian NN provides a better performance and is relatively more stable than the classical NN, it remains less stable than the Bayesian classifier. Thus, we can confirm that the Bayesian approach for ANNs improves the stability and performance of the conventional NN.

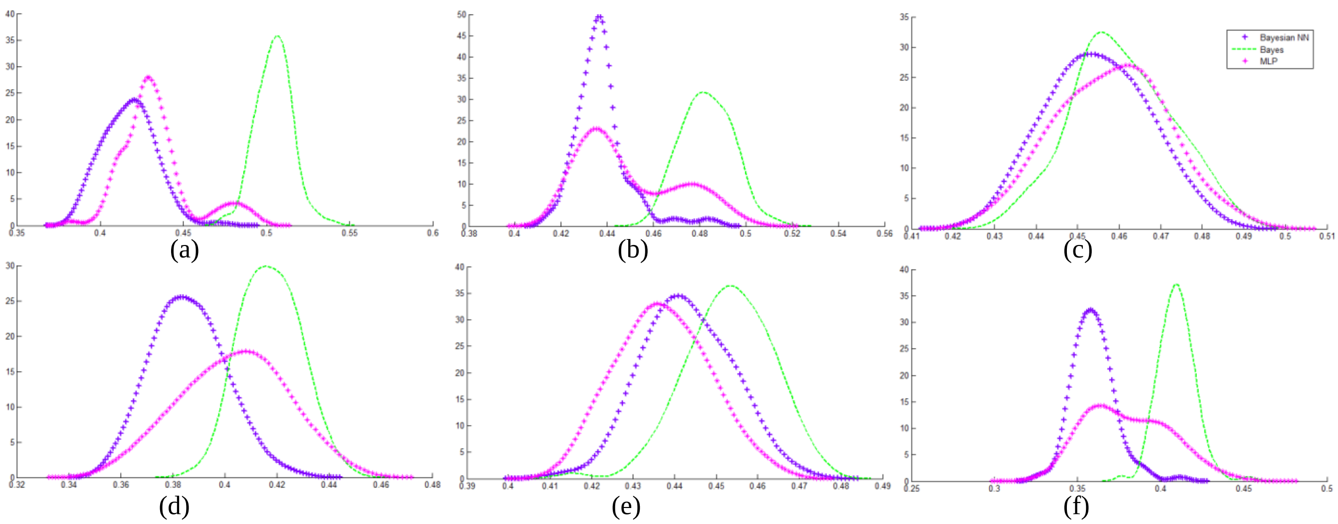


Figure 2: Error rate densities of Bayesian classifier (in green(**..**)), MLP (in pink(*)) and Bayesian NN (in purple(+)).

Table 2: Comparison results of ANN and Bayesian classifier for the Pearson system distributions.

Cases	**Distributions**										**Bayes**		**MLP**		**Bayesian NN**	
	Distribution 1					***Distribution 2***										
	Type	μ	σ^2	β_1	β_2	***Type***	μ	σ^2	β_1	β_2	***Mean***	***Variance***	***Mean***	***Variance***	***Mean***	***Variance***
a	8	40	100	0	3	3	60	100	1.26	4.9	0.5044	0.1154	0.4328	0.4460	0.4172	0.2274
b	1	10	100	0.63	3.28	3	30	100	1.26	4.9	0.4831	0.1151	0.4491	0.4501	0.4374	0.1369
c	8	40	100	0	3	2	60	100	0	2.34	0.4607	0.1403	0.4581	0.1687	0.4548	0.1414
d	6	40	100	1.32	5.35	2	60	100	0	2.34	0.4174	0.1213	0.4030	0.3912	0.3860	0.1950
e	4	40	100	1.71	7.3	6	60	100	1.32	5.35	0.4531	0.1043	0.4372	0.1162	0.4430	0.1077
f	4	40	100	1.71	7.3	2	60	100	0	2.34	0.4096	0.1184	0.3802	0.5564	0.3597	0.1643

ISBN: 978-1-61804-227-9

V. Application to Handwritten Digit Recognition

In this section, we study the handwritten digit recognition problem, which is still one of the most important topics in the automatic sorting of postal mails and checks' registration. The database used to train and test the different classifiers described in this paper was selected from the MNIST database. This database contains 60,000 training images and 10,000 test ones.

For the training and test sets, we select randomly, from the MNIST training and test sets respectively, single digit images from '0 'to '3'. Each class contains 1000 images for the both sets. Some images are shown in Fig.3.

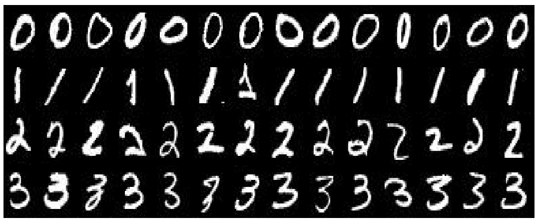


Figure 3: Sample images of MNIST database.

The most difficult step in handwritten digit recognition is to choose the suitable features. The chosen features must necessarily verify a non-exhaustive set of criteria such as stability, completeness, fast computation, powerful discrimination and invariance under the geometrical transformations. The invariant descriptors family proposed by Ghorbel in [6] satisfies the various criteria cited above. Thus, each image will be described by this type of invariants and Fourier descriptors (FD). The selected descriptors size is high (D = 14). In order to apply Bayesian rule, dimension reduction becomes necessary. The transformation matrices are estimated for both PCA and LDA methods from the training set, which transform the data to the appropriate dimensions subspace (two dimensions in our study). For the neural approach, we have used a MLP and a Bayesian NN with three layers having, respectively, 14, 10 and 4 neurons. In order to compare the classifiers' stability, we evaluate the classifiers' performance for 100 times using the k-folds cross validation algorithm (k=10 in our study). The misclassification rate (MCR) of each classifier is calculated on the test sets selected by the CV algorithm from the MNIST test set (N=400 images for each class). Figures 4.a and 4.b represent the MCR probability estimation using the four classifiers (PCA-Bayes, Fisher-Bayes MLP and Bayesian NN) for Fourier descriptors and Ghorbel descriptors, respectively.

In order to obtain meaningful comparison between the different types of classifiers, we evaluate their performances and stability degrees. Figure 4 shows the error rate probability densities estimated using the modified semi-bounded Plug-in algorithm. This algorithm is qualified by its sufficient precision on the stability aspects. In table 3, we summarize the MCR means and variances obtained for the two types of descriptors using the four classifiers. We note that these classifiers give the best results for Ghorbel descriptors. The MLP shows performance against the Bayesian classifier, but the superiority of its error rate variances proves that it is less stable than the statistical approaches. Although the Bayesian NN provides a better performance and is relatively more stable than the classical NN. Thus, we can confirm that the Bayesian approach for ANNs improves the stability and performance of the conventional NN.

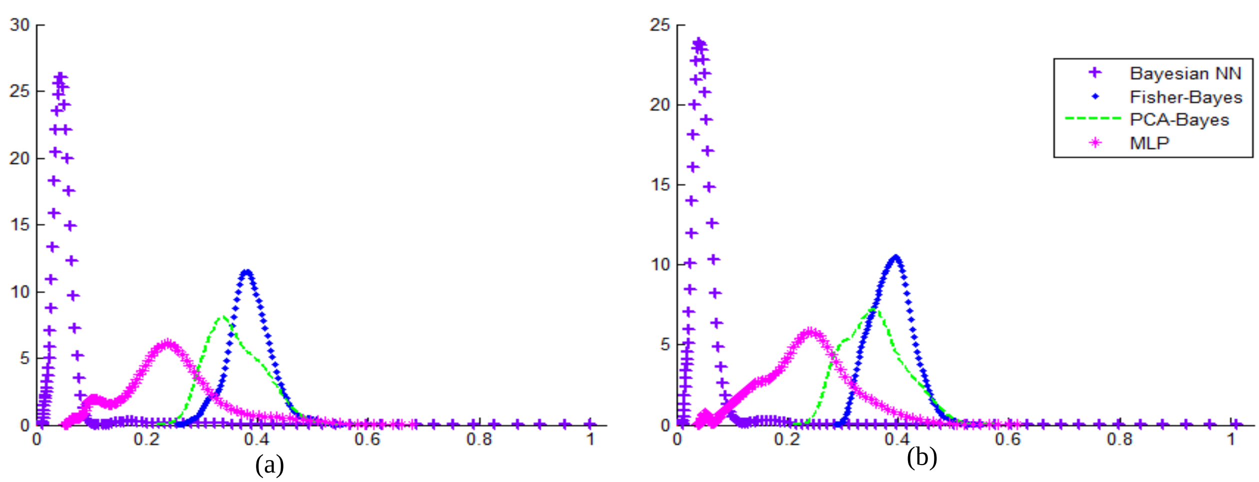


Figure 4: Error rate densities of PCA-Bayes (in green(--)), Fisher-Bayes (in blue(..)), MLP (in pink(*)) and Bayesian NN (in purple(+)) for Fourier descriptors (in the left) and Ghorbel descriptors (in the right).

Table 3: Comparison results of neural and statistical classifiers on the MNIST database.

	Fourier Descriptors		**Ghorbel Descriptors**	
	Mean	**Variance**	**Mean**	**Variance**
PCA-Bayes	0.3599	0.0022	0.3572	0.0026
Fisher-Bayes	0.3855	0.0014	0.3840	0.0012
MLP	0.2469	0.0072	0.2385	0.0057
Bayesian NN	0.0747	0.0037	0.0646	0.0011

VI. CONCLUSIONS

In this paper, a new criterion to comparing neural and Bayesian classifiers was proposed. In fact, a stability comparison based on estimating classifiers error rate probability density functions was presented.

The stochastic simulations demonstrated the superiority of the statistical approaches stability compared to the neural networks stability. In addition, the Bayesian approach for modeling NNs enhances their performance and stability. This study has provided a new conception to compare the stability results of the neural networks and other classifiers kinds. Another interesting point would be also to combine the classifiers to improve their stabilities.

REFERENCES

[1] A.M. Martinez, A.C. Kak, “PCA versus LDA”, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 23, no. 2, pp. 228-233, 2001.

[2] C. M. Bishop, “Neural networks for pattern recognition”. *Oxford university press*, 1995.

[3] D.J.C. MacKay, “A practical Bayesian framework for back-propagation networks”. *Neural Comput*, 4(3), 448–72, 1992.

[4] MACKAY, David JC, et al. Bayesian methods for neural networks: theory and applications. 1995.

[5] D.W. Miller, “Fitting frequency distributions”, *Book Resource*, second edition, 1998.

[6] F. Ghorbel, “Towards a unitary formulation for invariant image description: application to image coding”. *Annals of telecommunication*, vol. 53, France, 1998.

[7] F. Menasri, N. Vincent, E.Augustin, “Reconnaissance de chiffres farsi isolés par réseau de neurones à convolutions”, CIFED’08, pp. 127-132, 2008.

[8] G.P. Zhang, “Neural networks for classification: a survey”. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions*, vol. 30, no 4, p. 451-462, 2000.

[9] I.T. Jolliffe, “Principal Component Analysis”. *Springer-Verlag*, 1986.

[10] K. Fukunaga, “Introduction to Statistical Pattern Recognition”, second ed. *Academic Press*, 1990.

[11] K. Steven, Rogers, M. Kabrisky, “An Introduction to Biological and Artificial Neural Networks for Pattern Recognition”, *SPIE Optical Engineering Press*, vol. 4, 1991.

[12] K.Y. Tam, M.Y. Kiang, “Managerial applications of neural networks: The case of bank failure predictions”. *Management Science*, vol. 38, no. 7, pp. 926–947, 1992.

[13] L. K. Hansen and P. Salamon, “Neural network ensembles,” *IEEE Trans.Pattern Anal. Machine Intell.*, vol. 12, no. 10, pp. 993–1001, 1990.

[14] M. Paliwal, U.A. Kumar, “Neural networks and statistical techniques: A review of applications”, *Expert Syst. Appl.*, vol. 36, no. 1, pp. 2–17, 2009.

[15] M. P. Perrone and L. N. Cooper, “When networks disagree: Ensemble methods for hybrid neural networks”, *Neural Networks for Speech and Image Processing*, R. J. Mammone, Ed. London, U.K.: Chapman & Hall, 1993, pp. 126–142.

[16] M. Troudi, F. Ghorbel, “The generalised Plug-in algorithm for the diffeomorphism kernel estimate”. *International Conference on Systems, Control, Signal Processing and Informatics*, 2013.

[17] N.L. Johnson, S. Kotz, “Distribution in statistics: Continuous univariate distribution”, *Wiley-interscience*, vol. 1, New York, 1969.

[18] N. Morgan and H. Bourlard, “Generalization and parameter estimation in feedforward nets: Some experiments,” Adv. *Neural Inform. Process. Syst.*, vol. 2, pp. 630–637, 1990.

[19] R. Lepage, B. Solaiman, “Les réseaux de neurones artificiels et leurs applications en imagerie et en vision par ordinateur”. Montréal, 2003.

[20] S. Geman, E. Bienenstock, and T. Doursat, “Neural networks and the bias/variance dilemma”, *Neural Comput.*, vol. 5, pp. 1–58, 1992.

[21] S.M. Weiss and C.A. Kulilowski, “Computer Systems that Learn”. San Mateo, CA: Morgan Kaufmann, 1991.

[22] S. Saoudi, F. Ghorbel, A. Hillion, “Some statistical properties of the kernel-diffeomorphism estimator”, *Applied Stochastic Models and Data Analysis Journal*, Vol. 13, no. 1, pp. 39-58, 1997.

[23] S. Saoudi, F.Ghorbel, A.Hillion, “Nonparametric probability density function estimation on a bounded support: applications to shape classification and speech coding”, *Applied Stochastic Models and Data Analysis Journal*, vol. 10, no. 3, pp. 215–231, 1994.

[24] S. Saoudi, M. Troudi, F. Ghorbel, “An iterative soft bit error rate estimation of any digital communication systems using a non parametric probability density function”, *Eurasip Journal on wirless Communications and Networking*, 2009.

[25] U.A. Kumar, “Comparison of neural networks and regression analysis: A new insight”. *Expert Systems with Applications*, vol. 29, no. 2, pp. 424–430, 2005.

[26] W. Drira, F. Ghorbel, “Vers un classifieur de Bayes pour les grandes dimensions par analyse discriminante basée sur un estimateur non paramétrique de la L2-mesure de dépendance probabiliste”, *Récentes avancées en Reconnaissance de Formes Statistique*, ed. Art-pi, Tunis, www.arts-pi.org.tn, 2012.

[27] W. Patuwo, M.Y. Hu, M.S. Hung, “Two-group classification using neural networks”. *Decision Sciences*, vol. 24, pp. 825–845, 1993.

Study and Application of the Advanced Frequency Control Techniques H∞ in the Voltage Automatic Regulator of Powerful Synchronous Generators (Application under Gui/Matlab)

GHOURAF Djamel Eddine
Department of Electrical Engineering University of SBA
IRECOM Laboratory
BP 98 22000 Algeria
E-mail: jamelbel22@yahoo.fr

NACERI Abdellatif
Department of Electrical Engineering University of SBA
IRECOM Laboratory
BP 98 22000 Algeria

Abstract— This article presents a study of of the advanced frequency control techniques based on loop-shaping H∞ optimization method applied on automatic excitation control of powerful synchronous generators (AVR and PSS), to improve transient stability and its robustness of a single machine- infinite bus system (SMIB).The computer simulation results (static and dynamic stability), with test of robustness against machine parameters uncertainty (electric and mechanic), have proved that good dynamic performances, showing a stable system responses almost insensitive to large parameters variations, and more robustness using the robust H-infinity controller, in comparison with the classical Russian PID power system stabilizer . Our present study was performed using a GUI realized under MATLAB in our work.

Keywords— powerful synchronous generators and Excitations, AVR and PSS, , stability and robustness, GUI.

I. Introduction

Power system stability continues to be the subject of great interest for utility engineers and consumers alike and remains one of the most challenging problems facing the power community. Power system oscillations are damped by the introduction of a supplementary signal to the excitation system of a power system. This is done through a regulator called power system stabilizer. Classical PSS rely on mathematical models that evolve quasi-continuously as load conditions vary. This inadequacy is somewhat countered by the use of fuzzy logic in modeling of the power system. Fuzzy logic power system stabilizer is a technique of incorporating expert knowledge in designing a controller [1].

The Power System Stabilizer (PSS) is a device that improves the damping of generator electromechanical oscillations. Stabilizers have been employed on large generators for several decades; permitting utilities to improve stability constrained operating limits. The input signal of conventional PSS is filtered to provide phase lead at the electromechanical frequencies of interest (ie , 0.1 Hz to 5.0 Hz). The phase lead requirement is site-specific, and is required to compensate for phase lag introduced by the closed-loop voltage regulator.

The PSS conventional and the PSS control based on root locus and eigen value assignment design techniques have been widely used in power systems. Such PSS ensure optimal performance only at a nominal operating point and do not guarantee good performance over the entire range of the system operating conditions due to exogenous disturbances such as changes of load and fluctuations of the mechanical power. In practical power system networks, a priori information on these external disturbances is always in the form of a certain frequency band in which their energy is concentrated. Remarkable efforts have been devoted to design appropriate PSS with improved performance and robustness. These have led to a variety of design methods using optimal control [2] and adaptive control [3]. The shortcoming of these model-based control strategies is that uncertainties cannot be considered explicitly in the design stage. More recently, robust control theory has been introduced into PSS design which allows control system designers to deal more effectively with model uncertainties [4, 5, 6 and 7].H∞ based control approach is particularly appropriate for plants with unstructured uncertainty.

In this paper, a PSS based on H∞ robust control techniques is introduced and results are displayed in time response approach for studying stability of electric power system under different conditions.

Simulation results shown the evaluation of the proposed linear control methods based on this advanced frequency techniques applied in the automatic excitation regulator of powerful synchronous generators: the robust H∞ linear stabilizer and conventional PID control schemes against system variation in the SMIB power system, with a test of robustness against parametric uncertainties of the synchronous machines (electric and mechanic), and make a comparative study between these two control techniques for AVR – PSS systems.

II. Dynamic Power System Model:

2.1. Power System description

In this paper the dynamic model of an IEEE - standard of power system, namely, a single machine connected to an infinite bus system (SMIB) was considered [8]. It consists of a single synchronous generator (turbo-Alternator) connected through a parallel transmission line to a very large network approximated by an infinite bus as shown in figure 1.

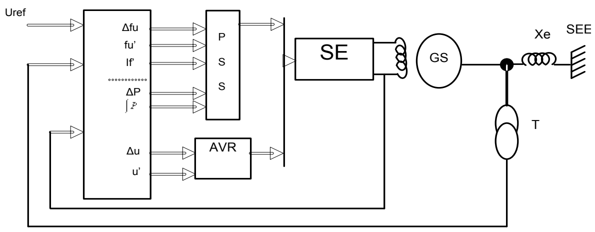


Fig. 1. Standard system IEEE type SMIB with excitation control of powerful synchronous generators

2.2. The permeances networks modeling (Park-Gariov) of powerful synchronous generators

In this paper we based on the permeances networks modeling of powerful synchronous generators for eliminating simplifying hypotheses and testing the control algorithm. The PSG model is defined by equations and Figure 2 and 3 below [8]:

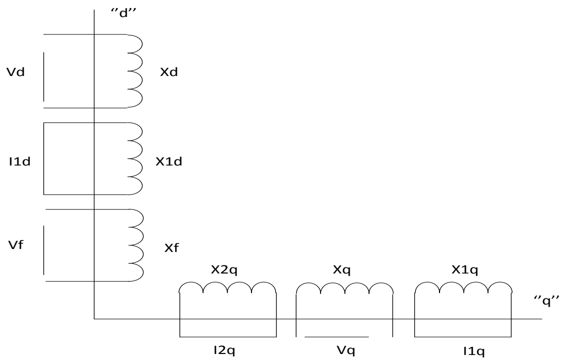


Fig. 2. PARK Transformation of the synchronous machine

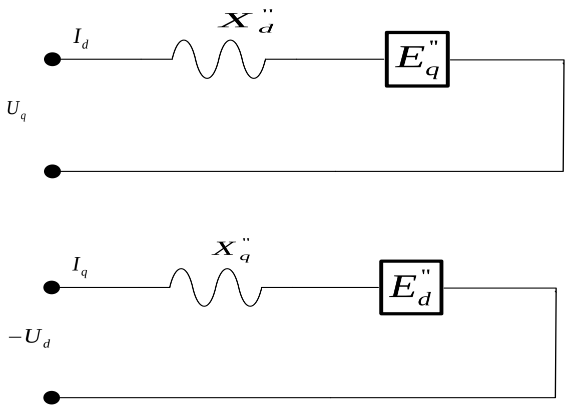


Fig.3. Equivalent diagrams simplifies of the synchronous machine with damping circuits (PARK-GARIOV model)

A. Currants equations:

$$I_q = (U_q - E_q^{''}) / X_d^{''} \qquad I_{1q} = (\Phi_{1q} - \Phi_{aq}) / X_{sr1q}$$
$$I_d = -(U_d - E_d^{''}) / X_q^{''} \qquad I_{2q} = (\Phi_{2q} - \Phi_{aq}) / X_{sr2q} \tag{1}$$
$$I_{1d} = (\Phi_{1d} - \Phi_{ad}) / X_{srd} \qquad I_f = (\Phi_f - \Phi_{ad}) / X_{sr}$$

$$E_q^{''} = \frac{\frac{1}{X_{sf}} \cdot \frac{X_f}{X_{ad}} E_q^{'} + \frac{1}{X_{sfd}} \cdot \frac{X_{fd}}{X_{ad}} E_{fq}^{'}}{\frac{1}{X_{ad}} + \frac{1}{X_{sf}} + \frac{1}{X_{sfd}}} \qquad E_d^{''} = \frac{\frac{1}{X_{sfq}} \cdot \frac{X_{fq}}{X_{aq}} E_{fd}^{'}}{\frac{1}{X_{ad}} + \frac{1}{X_{sfq}}} \tag{2}$$

B. Flow equations:

$$\Phi_{ad} = E_q^{''} + \left(X_d^{''} - X_s\right) I_d \; ; \; \Phi_{aq} = E_d^{''} + \left(X_q^{''} - X_s\right) I_q$$
$$\Phi_{1q} = \omega_s \int_0^{\Phi_{1q}} \left(- R_{1q} I_{1q}\right) dt \quad \Phi_{2q} = \omega_s \int_0^{\Phi_{2q}} \left(- R_{2q} I_{2q}\right) dt \tag{3}$$
$$\Phi_f = \omega_s \int_0^{\Phi_f} \left(- R_f I_f + U_{f0}\right) dt \quad \Phi_{1d} = \omega_s \int_0^{\Phi_{1d}} \left(- R_{1d} I_{1d}\right) .dt$$

C. Mechanical equations

$$d\delta = \left(\omega - \omega_s\right) dt \;,\; s = \frac{\omega - \omega_s}{\omega_s} \tag{4}$$

$$M_T + M_j + M_e = 0 \quad avec\ M_j : \text{moment d'inertie} \quad \left(M_j = -j \frac{d\omega}{dt}\right)$$
$$T_j \frac{d}{dt} s + \left(\Phi_{ad} . I_q - \Phi_{aq} . I_d\right) = M_T \quad \text{ou} \quad T_j \frac{d}{dt} s = M_T \text{ - } M_e \tag{5}$$
$$j \frac{d\omega}{dt} + \frac{P_e}{\omega_s} = M_T$$

2.3. Models of regulators AVR and PSS:

The AVR (Automatic Voltage Regulator), is a controller of the PSG voltage that acts to control this voltage, thought the exciter .Furthermore, the PSS was developed to absorb the generator output voltage oscillations [9].

In our study the synchronous machine is equipped by a voltage regulator model "IEEE" type – 5 [10, 11], as is shown in figure 4.

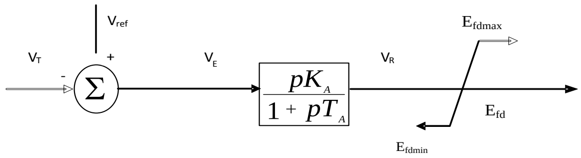


Fig. 4. A simplified " IEEE type-5" AVR

$$V_R = \frac{K_A V_E - V_R}{T_A} \quad , \quad V_E = V_{ref} - V_F \tag{6}$$

About the PSS, considerable's efforts were expended for the developpement of the system. The main function of a PSS is to modulate the SG excitation to [9, 12, 8].

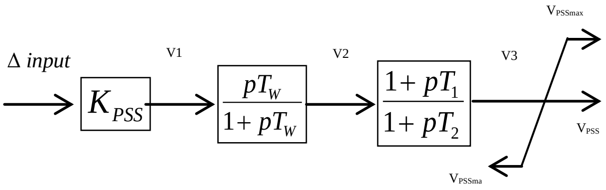


Fig.5. A functional diagram of the PSS used [8]

In this paper the PSS signal used, is given by: [13]

$$\begin{aligned}
&\dot{V}_1 = \frac{V_2 - V_1}{T_1} + \frac{T_2}{T_1}\dot{V}_2 \quad ; \\
&\dot{V}_2 = \frac{V_3 - V_2}{T_2} + \frac{T_3}{T_2}\dot{V}_2 \quad ; \qquad \Delta input = \begin{cases} \Delta P \ , \int p \\ or \\ \Delta\omega = \omega_{mach} - \omega_0 \\ and \\ \Delta I_f = I_f - I_{f0} \\ and \\ \Delta U_f = U_f - U_{f0} \end{cases} \\
&\dot{V}_3 = \frac{V_3}{T_W}\dot{V}_1 ; \ \dot{V}_1 = K_{PSS} \cdot \Delta input
\end{aligned} \tag{7}$$

2.4. Simplified model of system studied SMIB

We consider the system of figure 6. The synchronous machine is connected by a transmission line to infinite bus type SMIB. The transmission line with a resistance Re, and an inductance Le [8].

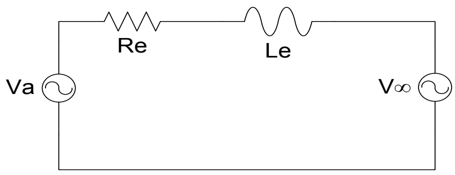


Fig.6. Synchronous machine connected to an infinite bus network

We define the following equation of SMIB system

$$V_{\infty odq} = P v_{\infty abc} = \sqrt{2}\, V_\infty \begin{bmatrix} 0 \\ -Sin\,(\delta - \alpha) \\ \cos\,(\delta - \alpha) \end{bmatrix} + L_e I'_{odq} + X_e \begin{bmatrix} 0 \\ -i_q \\ i_d \end{bmatrix} \tag{8}$$

III. The Robust PSS Based on Loop –Shaping H∞ Optimization

3.1. The H∞ theory:

Advanced control techniques have been proposed for stabilizing the voltage and frequency of power generation systems. These include output and state feedback control, variable structure and neural network control, fuzzy logic control [14,15, 16], Robust H_2 (linear quadratic Gaussian with KALMAN filter) and robust H_∞ control [17,18].

H_∞ approach is particularly appropriate for the stabilization of plants with unstructured uncertainty [18]. In which case the only information required in the initial design stage is an upper band on the magnitude of the modeling error. Whenever the disturbance lies in a particular frequency range but is otherwise unknown, then the well known LQG (Linear Quadratic Gaussian) method would require knowledge of the disturbance model [17]. However, H_∞ controller could be constructed through, the maximum gain of the frequency response characteristic without a need to approximate the disturbance model. The design of robust loop – shaping H_∞ controllers based on a polynomial system philosophy has been introduced by Kwakernaak and Grimbel [19, 20].

H_∞ synthesis is carried out in two phases. The first phase is the H_∞ formulation procedure. The robustness to modelling errors and weighting the appropriate input – output transfer functions reflects usually the performance requirements. The weights and the dynamic model of the power system are then augmented into an H_∞ standard plant. The second phase is the H_∞ solution. In this phase the standard plant is programmed by computer design software such as MATLAB [21-22], and then the weights are iteratively modified until an optimal controller that satisfies the H_∞ optimization problem is found [23].

Time response simulations are used to validate the results obtained and illustrate the dynamic system response to state disturbances. The effectiveness of such controllers is examined and compared with using the linear Robust H_∞ PSS at different operating conditions of power system study.

The advantages of the proposed linear robust controller are addresses stability and sensitivity, exact loop shaping, direct one-step procedure and close-loop always stable [17].

The H_∞ theory provides a direct, reliable procedure for synthesizing a controller which optimally satisfies singular value loop shaping specifications [24-23]. The standard setup of the control problem consist of finding a static or dynamic feedback controller such that the H-INFINITY norm (a uncertainty) of the closed loop transfer function is less than a given positive number under constraint that the closed loop system is internally stable.

The robust H_∞ synthesis is carried in two stages:

a. *Formulation*: Weighting the appropriate input – output transfer functions with proper weighting functions. This would provide robustness to modelling errors and achieve the performance requirements. The weights and the dynamic model of the system are hen augmented into H-INFINITY standard plant.
b. *Solution:* The weights are iteratively modified until an optimal controller that satisfies the H_∞ optimization problem is found.

Fig.7 shows the general setup of the problem design where: P(s): is the transfer function of the augmented plant (nominal Plant G(s) plus the weighting functions that reflect the design specifications and goals),

u_2: is the exogenous input vector; typically consists of command signals, disturbance, and measurement noises,
u_1: is the control signal, y_2: is the output to be controlled, its components typically being tracking errors, filtered actuator signals, y_1: is the measured output.

ISBN: 978-1-61804-227-9

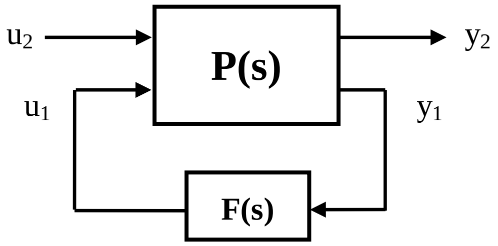


Fig.7. General setup of the loop-shaping H_∞ design

The objective is to design a controller F(s) for the augmented plant P(s) such that the input / output transfer characteristics from the external input vector u_2 to the external output vector y_2 is desirable. The H_∞ design problem can be formulated as finding a stabilizing feedback control law $u_1(s)$-$F(s).y_1(s)$ such that the norm of the closed loop transfer function is minimized.

In the power generation system including H_∞ controller, two feedback loops are designed; one for adjusting the terminal voltage and the other for regulating the system angular speed as shown on figure 8. The nominal system G(s) is augmented with weighting transfer function $W_1(s)$, $W_2(s)$, and $W_3(s)$ penalizing the error signals, control signals, and output signals respectively. The choice proper weighting functions are the essence of H_∞ control. A bad choice of weights will certainly lead to a system with poor performance and stability characteristics, and can even prevent the existence of solution to the H_∞ problem.

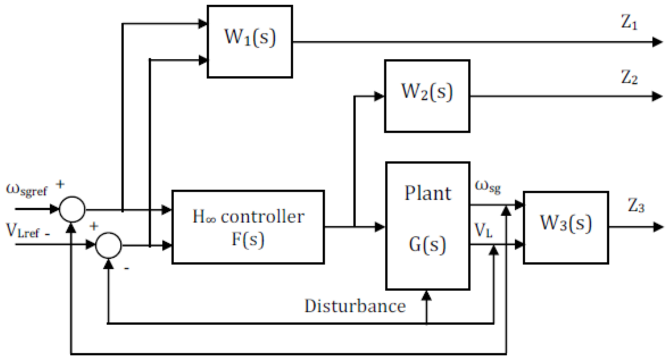


Fig.8. Simplified block diagram of the augmented plant including H_∞ controller

The control system design method by means of modern robust H-infinity algorithm is supposed to have some linear conventional PID test regulator.

It is possible to collect various optimal adjustment of such a regulator in different operating conditions into some database. Traditional Russian Power system stabilizer (realized on PID schemes) was used in this work as a test system, which enables to trade off regulation performance, robustness of control effort and to take into account process and measurement noise [17].

3. 2 GLOVER - DOYLE algorithm to synthesize a robust stabilizer PSS- H ∞

Problem solving of standard control is proposed as follows [22]:
1. Calculates the Standing regime established (RP) ;
2. Linearization of the control object (GS+PSS+AVR)
3. The main problem in H ∞ control and the definition of the control object increased P(s) in the state space:
   3-1. Choice of weighting functions: W_1, W_2, W_3
   3-2. The obtaining of the command object increased from weighting functions $W_{1,2,3}$.
4. Verify if all conditions to the ranks of matrices are satisfied, if not we change the structure of the weightings functions;
5. Choosing a value of γ (optimization level) ;
6. Solving two Riccati equations which defined by the two matrices H and J of HAMILTHON;
7. Reduction of the regulator order if necessary
8. By obtaining optimum values and two solutions of Riccati equations we get the structure of controller H ∞ and the roots of the closed loop with the robust controller;
9. We get the parameters of robust controller H_∞ in linear form 'LTI (SS state space, TF transfer function or ZPK zeros - pole - gains)
10. The simulation and realization of the stability study and robustness of electro-energy system under different functioning conditions.

The synthesis algorithm of the robust controller H∞ proposed in this work is clearly shown schematically by the flow chart of Figure 9

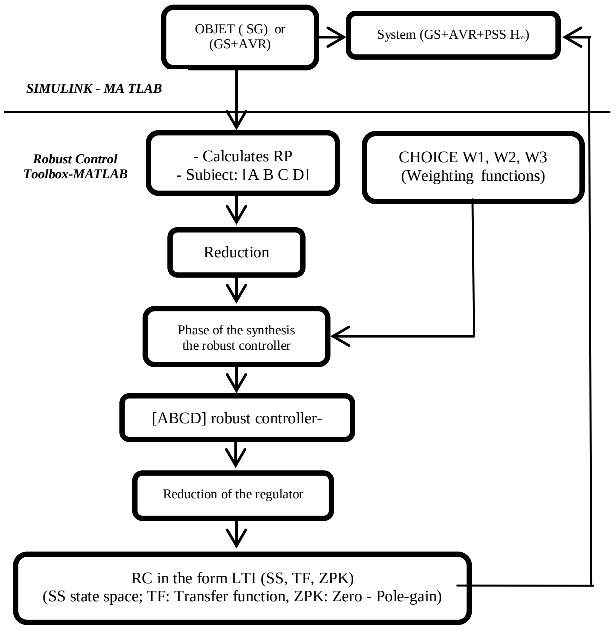


Fig.9 Synthesis algorithm robust controller of the excitation for a single machine

3.3 Structure of the power System withe Robust Controllers H∞

The basic structure of the control system a powerful synchronous generator with the robust controller shown in Figure 10 [3]

As command object we have synchronous generator with regulator AVR-FA (PID with conventional PSS), an excitation system (exciter) and an information block and measures (BIM) of output parameters to regulated.

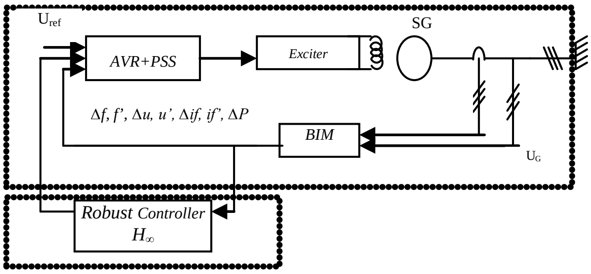


Fig 10. structure of the power system withe robust controllers H∞

IV . The Simulation Result Under GUI/ Matlab

A) Creation of a calculating code under MATLAB / SIMULINK

The "SMIB" system used in our study includes:

- A powerful synchronous generator (PSG) ;
- Tow voltage regulators: AVR and AVR-PSS connected to;
- A Power Infinite network line

We used for our simulation in this paper, the SMIB mathematical model based on permeances networks model culled Park-Gariov, and shown in Figure 11[13]

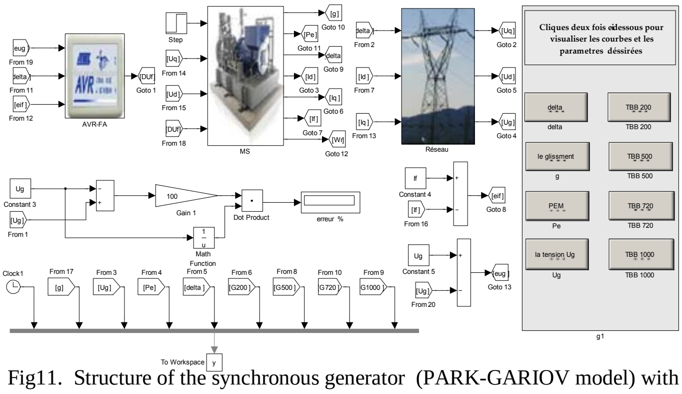


Fig11. Structure of the synchronous generator (PARK-GARIOV model) with the excitation controller under [13].

B) A Created GUI/MATLAB

To analyzed and visualized the different dynamic behaviors we have creating and developing a "GUI" (Graphical User Interfaces) under MATLAB .This GUI allows as to:

- Perform control system from PSS and H∞ -PSS controller;
- View the system regulation results and simulation;
- Calculate the system dynamic parameters ;
- Test the system stability and robustness;
- Study the different operating regime (under-excited, rated and over excited regime).

The different operations are performed from GUI realized under MATLAB and shown in Figure 12.

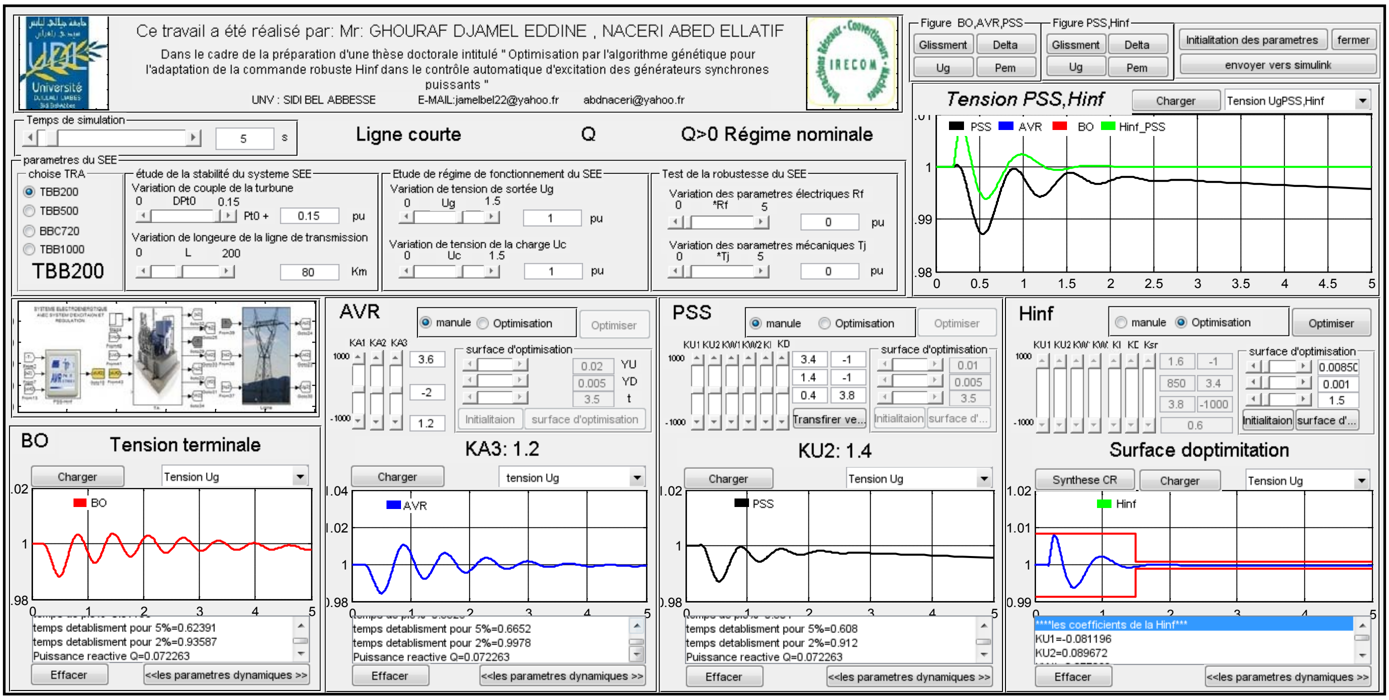


Fig.12.The realised GUI / MATLAB

ISBN: 978-1-61804-227-9

C) *Simulation result and discussion*

- ***Stability study***

We performed an perturbations by abrupt variation of turbine torque ΔTm of 15% at t = 0.2s,

The following results (Table 1 and figure 13, 14 and 15) were obtained by studying the "SMIB" static and dynamic performances in the following cases:

1. SMIB in open loop (without regulation) (OL)
2. Closed Loop System with the conventional stabilizer PSS-FA and robust control H∞ -PSS [13].

We simulated three operating: the under-excited, the rated and the over-excited.

Our study is interested in the powerful synchronous Generators of type: TBB-200, TBB-500 BBC-720, TBB-1000 (parameters in Appendix) [13].

Table 1 presents the TBB-1000 static and dynamic performances results in (OL) and (CL) with PSS and H∞ -PSS, for an average line (Xe = 0.3 pu), and an active power P=0.9 p.u , for more details about the calculating parameters see GUI-MATLAB in the Appendix created.

Where: α: Damping coefficient ε %: the static error, d%: the maximum overshoot, t_s: the setting time

TABLE 1 THE "SMIB "STATIC AND DYNAMIC PERFORMANCES

Damping coefficient α				The static error		
Q	**OL**	**PSS**	**H∞ -PSS**	**OL**	**PSS**	**H∞ -PSS**
-0.1372	Unstable	-1.956	-3.767	Unstable	1.459	negligible
-0.4571	Unstable	-1.926	-3.659	Unstable	1.461	negligible
0.1896	-0.2061	-1.966	-3.876	-5.933	1.386	negligible
0.3908	-0.2245	-1.850	-3.769	-5.802	1.170	negligible
0.5078	-0.3577	-1.412	-3.211	-4.903	0.659	negligible
0.6356	-0.3660	-1.401	-3.109	-4.597	0.683	negligible
The setting time for 5%				**The maximum overshoot %**		
Q	**OL**	**PSS**	**H∞ -PSS**	**OL**	**PSS**	**H∞ -PSS**
-0.1372	Unstable	1,534	rapid	9.458	8,766	3.768
-0.4571	Unstable	1,558	rapid	9.254	8,632	3.548
0.1896	-	1,526	rapid	9,249	8,811	4.691
0.3908	-	1,621	rapid	9,075	8,292	4.773
0.5078	8,387	2,125	rapid	8,053	6,328	3.729
0.6356	8,197	2,141	rapid	7,426	6,279	3,612

In the Figures 13,14 and 15 show an example of simulation result with respectively: 's' variable speed , 'delta' The internal angle, 'Pe' the electromagnetic power system, 'Ug' the stator terminal voltage; for powerful synchronous generators TBB-1000 with P = 0.9, Xe = 0.3, Q1 = -0.1372 (pu)

- ***robustness tests***

In a first step we performed an variations electrical parametric (increase 100% of R). then, we performed an variations mechanical parametric (lower bound 50% of inertia J)

The simulation time is evaluated at 8 seconds.

We present in Figure 13 (For electrical uncertainties) and Figure 14 (For mechanical uncertainties)

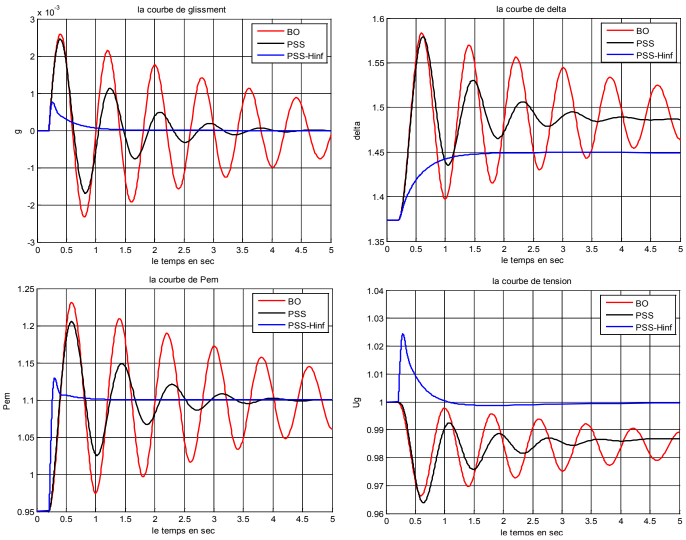


Fig .13. functioning system in the under-excited used of TBB-1000 connected to a average line with PSS , H∞ -PSS and OL (stability study)

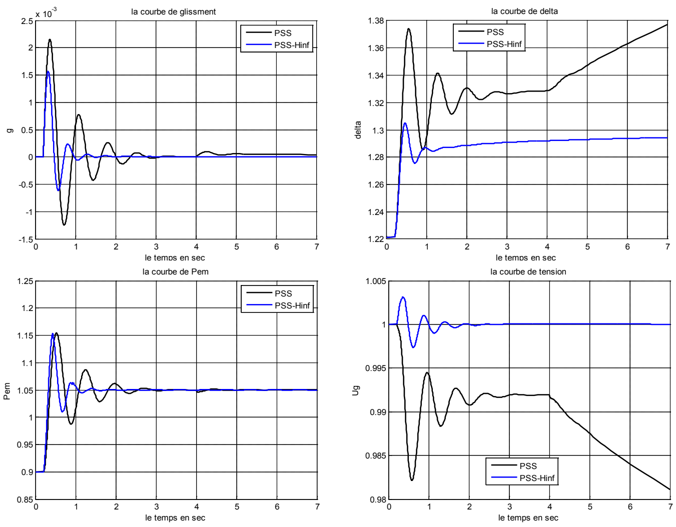


Fig .14.functioning system in the under-excited used of TBB-1000 connected to a average line with PSS , H∞ -PSS and OL (robustness tests)

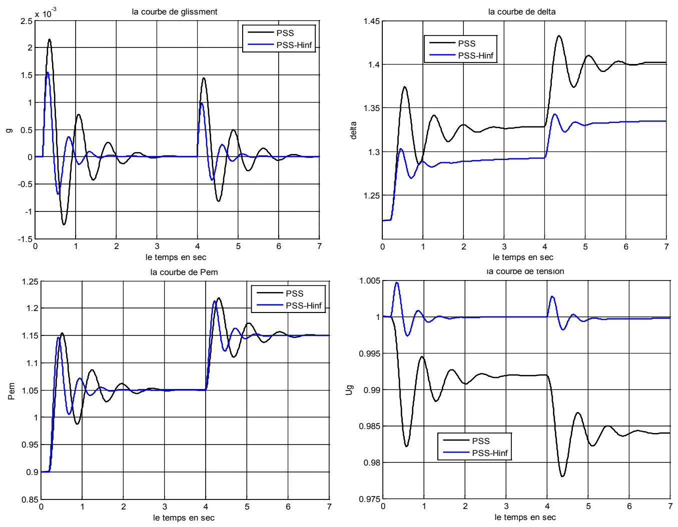


Fig .15. functioning system in the under-excited used of TBB-1000 connected to a average line with PSS , H∞ -PSS and OL (robustness tests)

From the simulation results, it can be observed that the use of H∞ -PSS improves considerably the dynamic performances (static errors negligible so better precision, and very short setting time so very fast system, and we found that after few oscillations, the system returns to its equilibrium state even in critical situations (specially the under-excited regime) and granted the stability and the robustness of the studied system.

V. CONCLUSION

This paper proposes an advanced control method based on advanced frequency techniques: a Robust H∞ Power system stabilizer based on loop-shaping optimization methods applied on the system AVR - PSS of powerful synchronous generators, to improve transient stability and its robustness of a single machine- infinite bus system (SMIB). This concept allows accurately and reliably carrying out transient stability study of power system and its controllers for voltage and speeding stability analyses. It considerably increases the power transfer level via the improvement of the transient stability limit.

The computer simulation results (with test of robustness against electric and mechanic machine parameters uncertainty), have proved a high efficiency and more robustness with the Robust H∞ PSS, in comparison using a Conventional Test stabilizer (with a strong action) realized on PID schemes, showing stable system responses almost insensitive under different modes of the station. This robust loop shaping H∞ generator voltage controller has the capability to improve its performance over time by interaction with its environment.

REFERENCES

[1] Ali.Z.M, and ELsherbiny M.," effect of both mechanical and excitation system on power system and methods of improvement.", M.sc, Egypt, Assuit university,2003.

[2] ABIDO, M. A.—ABDEL-MADJID, Y.L.: Coordinated Design of PSS and SVC Based Controller to Enhance Power System Stability, Electrical Power and Energy Systems 25 (2003), 695-704.

[3] GUPTA, R.—BANDYO, B.—KULKARNI, A. M.: Design of Power System Stability forSingle Machine System Using Robust Fast Output Sampling Feedback Technique, Electrical Power Systems Research 65 (2003), 247-257.

[4] BOUHAMIDA, M.—DENAI, M. A.—MOKHTARI, A.—BOU-HENNA, A.: GH∞ Robust Power System Stabilizer Design with Parametric Uncertainty, International Conference on Electrical and Electronics Engineering (ELECO'99), 1-5 December 1999, Bursa, Turkey.

[5] ZIAD M. M. ALI , ALEXANDER I.: ROBUST TECHNIQUES FOR DESIGNING POWER SYSTEM STABILIZER Journal of Theoretical and Applied Information Technology 2009

[6] CHEN, S.—MALIK, O. P.: H∞ Optimisation- Based Power System Stabiliser Design, IEE Proc. Gener. Trans. Distr. 142 No. 2 (1995), 179-184.

[7] FOLLY, K.—NAOTO, Y.—HIRISHI, S.: Design of Hinf -PSS Using Numerator- Denominator Uncertainty Representation, IEEE Trans. on Energy Conversion, 12 No. 1 (1997), 45-50.

[8] S. V. SMOLOVIK "Méthodes de modélisation mathématique des processus transitoires des générateurs synchrones plus usuels et non traditionnels dans les systèmes électro -énergétiques" Thèse doctorat d'état, Institut polytechnique de Leningrad, 1988 (traduit du Russe).

[9] LA. GROUZDEV, A.A. STARODEBSEV, S.M. OUSTINOV "Conditions d'application des meilleurs amortissements des processus transitoires dans les systèmes énergétiques avec optimisation numérique des paramètres du régulateur AVR-FA" Energie -1990-N°ll-pp.21-25 (traduit du russe).

[10] P.M. ANDERSON, A. A. FOUAD "Power System control and Stability", IEE Press, 1991.

[11] Hong Y.Y. and Wu W.C., « A new approach using optimization for tuning parameters of power system stabilizers », IEEE Transactions on Energy Conversion, vol. 14, n°. 3, pp. 780–786, Sept. 1999.

[12] DeMello F.P., Flannett L.N. and Undrill J.M., « Practical approach to supplementary stabilizing from accelerating power », IEEE Trans., vol. PAS-97, pp, 1515-1522, 1978.

[13] GHOURAF D.E., "Study and Application of the advanced frequency control techniques in the voltage automatic regulator of Synchronous machines', Magister Thesis, UDL-SBA, 2010 (In French).

[14] Kwakernaak H., Sivan R. Linear Optimal Control Systems, Wiley-Interscience, 1972.

[15] NACERI A., SMOLOVIK S.V. 'Estimation des lois de commande robustes H2 et Hinf dans les contrôleurs automatiques d'excitations des générateurs synchrones puissants' // Proc. VI conférence internationale ''problèmes scientifiques et techniques'', St. Petersbourg, 6 – 7 Juin 2002 (Tom. 1), édition SPBSPU, pp. 222 -228 (in Russain).

[16] Felix R. A., Sanchez E.N, Loukianov A.G., "Synchronous generator control combining sliding modes and neural networks" . American Control Conference, Vol. 5, Issue 4-6, June 2003, No. 4, pp. 4065-4070.

[17] Kwakernaak H. "Robust control and H∞ optimizatio n, Tutorial paper", Automatica, vol. 29, № 2. June 1994, pp. 201-207.

[18] A. Anderson and. A. Fouad ,"Power System Control and stability". IEEE Power systems,1994.

[19] MATLAB Math library User's Guide, by the Math works Inc., 2006.

[20] R. Y. Chiang, M. G. Sofonov "Robust control Toolbox User's guid", the Mathworks, INC, 1992.

[21] M. Sofonov , D. J. Limbeer, R. Y. Chiang "Simplifing the H∞ PSS theory via loop shaping matrix pencil and descriptor concept ", INT. J. Of control, vol. 50, № 6.1994, pp 2467-2488.

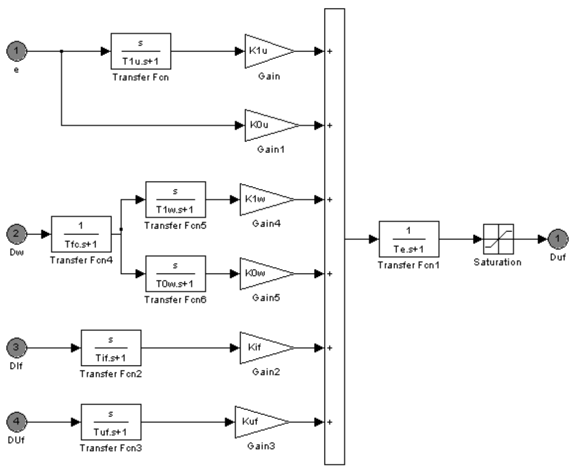
[22] K. Glover, J.C. Doyle, P.P. Khargonekar, B.A. Francis "State-space solutions to standard H2 and H∞ control problems", IEEE Trans. On A.C. 19 89, vol.34, № 8, pp.834-847.

[23] Grimbel M.J., "New direction in adaptive control theory and application". IFAC – IFIP- IMACS Conf., Control of industrial systems, Belfort, France 1997, pp. 95 – 110.

[24] Naceri, A. N Belaev, S. V Smolovik "rational Choices of the weight functions in the process of construction of the robust H∞ regulators for synchronous generators", the IX international conference "High intellectual technology of sciences", SaintPetersburgg, January - February 2002, pp. 251-257.

APPENDIX

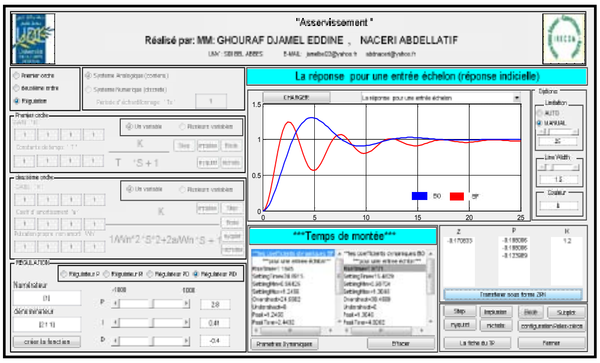
1. The PSS-AVR model



2. Parameters of the used Turbo –Alternators

Parameters	***TBB-200***	***TBB-500***	***BBC-720***	***TBB1000***	**Units of measure**
power nominal	200	500	720	1000	MW
Factor of power nominal.	0.85	0.85	0.85	0.9	p.u.
X_d	2.56	1.869	2.67	2.35	p.u.
X_q	2.56	1.5	2.535	2.24	p.u.
X_s	0.222	0.194	0.22	0.32	p.u.
X_f	2.458	1.79	2.587	2.173	p.u.
X_{sf}	0.12	0.115	0.137	0.143	p.u.
X_{sfd}	0.0996	0.063	0.1114	0.148	p.u.
X_{sf1q}	0.131	0.0407	0.944	0.263	p.u.
X_{sf2q}	0.9415	0.0407	0.104	0.104	p.u.
R_a	0.0055	0.0055	0.0055	0.005	p.u.
R_f	0.000844	0.000844	0.00176	0.00132	p.u.
R_{1d}	0.0481	0.0481	0.003688	0.002	p.u.
R_{1q}	0.061	0.061	0.00277	0.023	p.u.
R_{2q}	0.115	0.115	0.00277	0.023	p.u.

3. Dynamics parameters calculated through GUI-MATLAB



Using Intermediate Data of Map Reduce for Faster Execution

Shah Pratik Prakash, Pattabiraman. V

***Abstract*—**Data of any kind structured, unstructured or semi-structured is generated in large quantity around the globe in various domains. These datasets are stored on multiple nodes in a cluster. MapReduce framework has emerged as the most efficient technique and easy to use for parallel processing of distributed data. This paper proposes a new methodology for mapreduce framework workflow. The proposed methodology provides a way to process raw data in such a way that it requires less processing time to generate the required result. The methodology stores intermediate data which is generated between map and reduce phase and re-used as input to mapreduce. The paper presents methodology which focuses on improving the data reusability, scalability and efficiency of the mapreduce framework for large data analysis. MongoDB 2.4.2 is used to demonstrate the experimental work to show how we can store and reuse intermediate data as a part of mapreduce to improve the processing of large datasets.

***Keywords*—**MapReduce, Intermediate data management, MongoDB, Architecture aware mining.

I. Introduction

COMPANIES across various industries have been archiving data in their own required format and sizes. Retrieving Statistical, Analytical and other various forms information require high amount of processing of large size raw data, which is a costly task in terms of resource and time [7]. Present cloud based product having data distributed across several nodes uses MapReduce [4] framework for the processing of data with a certain level of optimization.

MapReduce is a parallel programming model which is very efficient for processing large data set across multiple core machine or nodes of a cluster. It is customizable so that one can implement their own logic for information retrieval. The workflow developed upon MapReduce achieves higher performance than traditional solutions. MapReduce model is based on two primary process map (Mapper) and reduce (Reducer), accompanied by combiner and finalize for performing reduction at map phase and operation on the result of reduce phase respectively. Analysis of raw data is carried out by workflow, which consists of several mapreduce jobs. The map task/phase is performed on either the raw data or on the result of previous mapreduce jobs. This phase performs calculations defined by the user and generates a key value pair as an intermediate data. The key and value represent a relation similar to a hash map in a programming construct. The map phase can emit zero or more times and same key with different values.

Map:

Raw data / Result of map reduce job ⟶ *{key, value}*

The result of mapping phases is the intermediate data which are then passed to reduce phase. The intermediate data is a list of values corresponding to unique key.

Intermediate data:

{key1, value1}
{key2, value2}
{key1, value3}
{key2, value4}

⇒

{key1, [value1, value3]}
{key2, [value2, value4]}

On reduce phase the user defined reduce function is performed on intermediate data. This can be ranging from sum, average, standard deviation to predictive analysis of values corresponding to the unique key. The result of reduce phase generates the same or different type of key with the result of the processed value list.

Reduce:

{key, [value1, vaue2]} ⟶ *{key3, value}*

In mapreduce framework the intermediate data generated is being deleted currently, so any mapreduce has to perform operation on raw data frequently. This paper focuses on storing and re-using the intermediate data and performing mapreduce more efficiently to generate the required output.

This was an introductory section to mapreduce framework describing its general execution and elements. Section II introduces how intermediate data are handled in MongoDB. Section III presents the proposed methodology. Section IV shows results of existing and proposed methodology. Section V is the analysis of experiment comparing existing system and proposed methodology. Section VI concludes this paper and present brief introduction to future work to be carried out. Lastly, we provide the list of references.

II. INTERMEDIATE DATA IN MAPREDUCE FRAMEWORK

Large dataset requires multilevel and multipath mechanism for computation. Workflow is formed composed of number of map reduce task. In a workflow the data generated by one mapreduce job is input to another and so on. It is found that several mapreduce jobs of different workflow are repeated over the raw data. This kind of execution is found in analysis, logging, transaction, image processing and other different types of complex task. As the mapreduce jobs are repeated, same intermediate data is generated which are either deleted or stored but not used. The intermediate data is a set for key value pair where key is unique and the value corresponding to the key is a list of values emitted by the map phase. The intermediate data is lost if its size is limited to temporary storage else is written to disk [2]. This intermediate data is then passed to reduce phase. Further, in this section follows MongoDB architecture and intermediate data in the mapreduce framework in MongoDB.

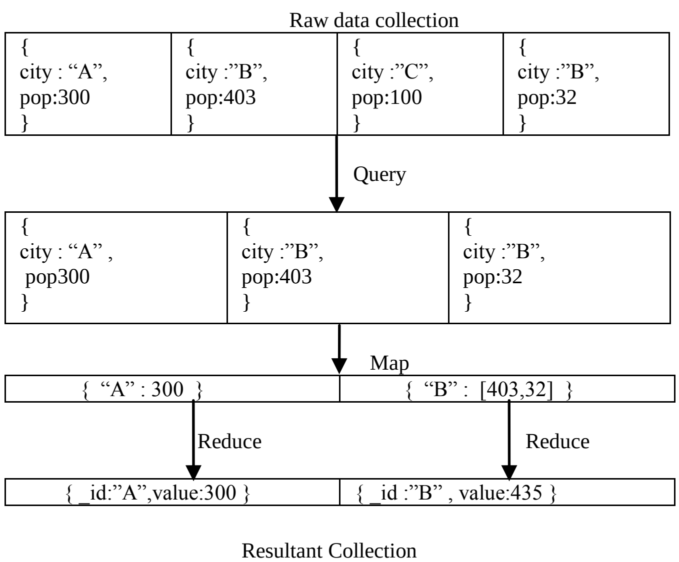


Fig 1. MapReduce operation in MongoDB.

A. MongoDB and MapReduce Framework

MongoDB is a NoSQL document-oriented database system [9]. MongoDB uses a memory map file that directly map disk data file to a memory byte array where data accesses and processing is carried out. In MongoDB "*database*" are represented in a similar pattern as RDBMS, incorporating various other objects like tables, indexes, triggers etc. Tables in MongoDB are referred as "*collection*", each can have zero or more record which is referred as "*document*". Columns in RDBMS have "*field*" as their counterpart in MongoDB.

MapReduce in MongoDB increases performance and utilization of resource [8]. Figure 1 shows a complete mapreduce operation in mongodb. It is a two-step process, map and reduce. On submission of user defined functions map and reduce to mapReduce, firstly map task is performed followed by reduce task. The map task processes either every document of the given collection or set of documents based on the specified query result. The map function emits key and its related value obtained during processing, this is at document level so there will be bunch of key-value pair, with the possibilities of pair having same named key with different values. These emitted pairs are stored in in-memory as temporal data and on memory size overflow they are stored on local disk. Figure 2 shows the existing mechanism of map phase. In reduce phase the reduce function performs a remote procedure call to mapper nodes to gather the intermediate data which is a pair of key and corresponding list of values. After the values are transferred from mapper node the user defined code processes these values for aggregation, statistical or analysis and the result is stored on a specified output collection. Figure 3 shows the existing mechanism of reduce phase.

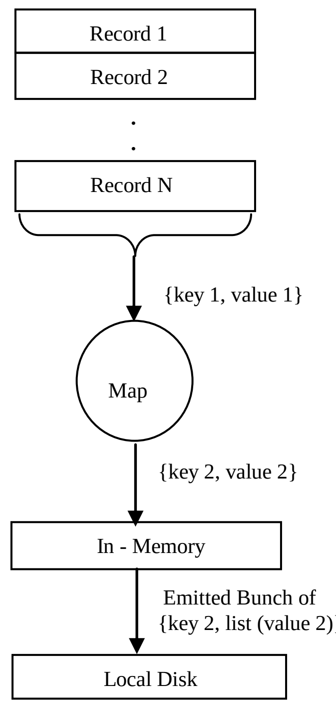


Fig 2. Existing Map phase

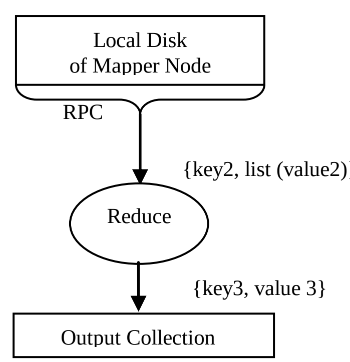


Fig 3. Existing Reduce Phase

B. Intermediate Data in MongoDB

The intermediate data are generated between the map and reduce phase. In mapreduce the reduce function is performed only after all map tasks of job are performed. At this point the data emitted by map is sent to the reducer. During this transfer the data from different nodes are collected and values of identical key are appended to list corresponding to the key, having the form:

{key, list(value)}

This pairs are stored in-memory and on overflow are placed on disk. They are then sorted based on key and sent to corresponding reducers [3].

C. Limitations

The mapreduce framework defines a job, a complex analytical task is performed as a workflow which is composed of number of mapreduce jobs. On daily basis several workflows are initiated in a company. Some of the jobs in are common among workflows. This leads to processing of same raw data again among different workflow.

The proposed methodology focuses on following areas of improvement in mapreduce framework:

1) *Input Size:* Processing large number of records.

2) *Processing Time:* Complex processing of data for emitting key value pair in map phase increases the overall time of workflow.

3) *Intermediate Data:* Generated intermediate data left unused on disk.

III. PROPOSED METHODOLOGY

The proposed methodology addresses the limitations described in section II. Performing complex calculations on raw data requires some mapreduce jobs to be repeated at some point in workflows affecting overall execution time, the proposed methodology uses a comparison of stored required pattern [6] with the minimal set of records of the stored intermediate data collection, which on satisfying the condition alters the map phase for the actual reduce phase to reduce the processing as well as the size of inputs. Thus a decision mechanism is established at the start of mapreduce job to decide which set of data and map task to use as an input to the job [2].

Figure 4 and 5 shows the conceptual view of the proposed methodology. Using this methodology requires less processing with raw dataset, only on first computation and updating of the dataset we perform mapreduce directly on raw dataset. On the direct computation on the raw dataset we store [3] or merge the intermediate dataset in a separate collection which are likely to be used frequently in the future as an input pattern to reduce phase and also compute the reduce task.

This methodology is useful for repetitive jobs. The map task and collection on which it is to be performed jobs are decided based on processing to be carried out over the sample records of the raw data. This is like a pattern recognition process [5]. This is carried out by applying mapreduce with reduce phase just returning the emitted elements of the map phase. The output is compared with the required pattern. On success we alter the map function implantation and collection (collection having intermediate data) which provides same data as we would have found using mapreduce on raw data collection. Thus the input changes to intermediate collection and implementation of map phase just becomes re emitting the key and values from list corresponding to the key which requires much less processing and the reduce task receives the same input so output as required receives the same input so output as required.

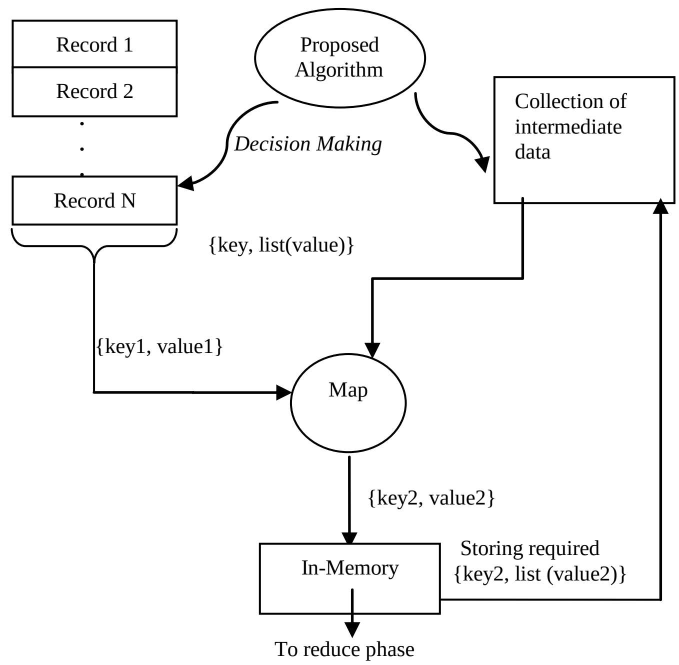


Fig 4. Proposed methodology for map phase

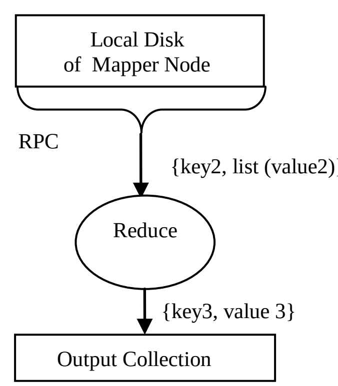


Fig 5. Reduce Phase of proposed system.

ISBN: 978-1-61804-227-9

Algorithm 1 Store intermediate data

```
/* map function*/
m <= function () {
/*
  In – between calculation
*/
     emit (key, this.value);
}
/* reduce function*/
r <= function (c, p) {
   return p;
}
db.rawdatacollection.mapReduce (m, r
  {
    out: intermediate data collection
  }
);
```

Algorithm 1 shows the method how to store intermediate data using mapreduce. The method comprises of map and reduces functions where the map function can have implementation focusing on that results that are commonly generated in the form of intermediate data in mapreduce workflow. The reduce function here simply returns the list of values captured as input for a key. MapReduce is performed on collection, storing raw data and the result is stored in a separate collection for intermediate dataset.

Algorithm 2 Decision making on raw data

```
m <= user defined map function
r   <= user defined reduce function
db.rawdatacollection.mapReduce (m, r,
   {
          out: scratch collection,
          query: select sample records
   }
);
If (output: result satisfies the required pattern)
        /* alter map function and perform mapreduce on
        intermediate data collection*/
          newM <= (function processing intermediate data
          emit key value pair)
          db.intermediatecollection.mapReduce (newM, r
        {
      out: resultant collection
        }
          ) ;
else
/* perform mapreduce on raw data collection directly*/
db.rawdatacollection.mapReduce (m, r,
       {
   out: resultant collection
       }
);
```

mapReduce syntax in MongoDB:

```
db.collectionName.mapReduce (
mapFunction,
 reduceFunction,
   {
       out: {
                         <action>: <collectionName>
        [, db: <dbName>]
        [, sharded: <boolean>]
        [, nonAtomic: <boolean>]
      },
       query: // clause to restrict records,
       jsMode: <boolean>,
       ...
          }
);
```

Algorithm 2 is the proposed method used for deciding which collection and mapping task to be applied for a mapreduce job. Initially map and reduce functions are defined as per user requirement, and then tested on very small number of documents of collection having the raw data. To restrict the sample size query parameter of mapReduce function can be used. This concept is similar to scratch database or temporary table in RDBMS. The output of this job is compared with the pattern which is the way how documents of intermediate data collection would be stored. If the pattern matches the document result of scratch collection we alter the collection and map function on which the mapreduce is to be performed. The mapreduce will now be executed on intermediate data collection and the implementation of map is changed such that every value of the list corresponding to the key is emitted as a key value pair. The reduce function is not changed so the output result and the collection where the result is to be stored remains the same. This effect of this is a very small operation is required to process the data at map phase which was taking comparatively more time for complex processing. Moreover the number of inputs document for job is less while intermediate collection is considered compared to raw data collection. New intermediate documents also merge or update with documents collection periodically for consistency. This is how the overall processing is improved leading to effective workflow.

IV. RESULT OF EXPERIMENT

This section addresses the experimental setup and the experiment performed using the existing and proposed methodology.

A. *Experimental Setup*

The experiment was performed using MongoDB 2.4.2 on 2 GB RAM node. NetBeans IDE 7.1.2 was used for executing java application. The raw data was collection of documents providing city wise population for different states. Figure 6 shows sample documents from the collection. 20 lakh documents were initially used for evaluation. The job is to find total state wise population. Further are the results of the experiments carried out.

{"city": "LOWER KALSKAG", "loc": [-160.359966, 61.51377], "pop": 291, "state": "AK", "_id": "99626"}
{"city": "MC GRATH", "loc": [-155.585153, 62.967153], "pop": 618, "state": "AK", "_id": "99627"}
{"city": "MANOKOTAK", "loc": [-158.989699, 59.009559], "pop": 385, "state": "AK", "_id": "99628"}
{"city": "FLAGSTAFF", "loc": [-111.574109, 35.225736], "pop": 26878, "state": "AZ", "_id": "86004"}
{"city": "EUFAULA", "loc": [-85.165605, 31.905063], "pop": 14189, "state": "AL", "_id": "36027"}
{"city": "DOZIER", "loc": [-86.366315, 31.506614], "pop": 741, "state": "AL", "_id": "36028"}
{"city": "PERRY", "loc": [-92.787976, 35.042732], "pop": 648, "state": "AR", "_id": "72125"}
{"city": "PERRYVILLE", "loc": [-92.847171, 34.970096], "pop": 3851, "state": "AR", "_id": "72126"}
{"city": "PLUMERVILLE", "loc": [-92.620435, 35.157466], "pop": 1940, "state": "AR", "_id": "72127"}

Fig 6. Raw data documents (records) from the collection

Firstly, the mapreduce was executed on the raw data collection directly. The execution input size and time size calculated is presented in Figure 7. Same required job was performed using the proposed methodology where first the intermediate data is stored in separate collection and then before performing the required job check is made that if the pattern of intermediate data exists if true, we modify the map implementation with reduce implementation same as that used in existing systems and calculate the input size and time taken. Figure 8 shows the proposed methodology result.

The time taken for retrieving sample results from the intermediate data collection and comparing is 250ms. Thus, combining processing time of mapreduce using intermediate data and time for pattern check is 10969ms which is less than processing time of existing methodology. Input size is half of the size used in existing methodology.

```
{
    "result" : "result",
    "timeMillis" : 39056,
    "counts" : {
        "input" : 2000000,
        "emit" : 2000000,
        "reduce" : 100000,
        "output" : 100000
    },
    "ok" : 1,
}
```

Fig 7. Input size and processing time in existing system.

```
{
    "result" : "Intresult",
    "timeMillis" : 10719,
    "counts" : {
        "input" : 100000,
        "emit" : 2000000,
        "reduce" : 100000,
        "output" : 100000
    },
    "ok" : 1,
}
```

Fig 8. Input size and processing time in proposed methodology.

V. ANALYSIS OF EXISTING SYSTEM AND PROPOSED METHODOLOGY.

The evaluation of proposed methodology was performed taking the aggregation experiment. This experiment aim is to compare the performance of existing system and proposed model, calculating output with new intermediate dataset and reducing the size of input documents for generating the same output as using the raw data documents.

A. Intermediate Data and Result in Existing System

Firstly mapreduce was carried out on the documents of raw data collection directly and output was stored on resultant collection. Repeatedly performing the task for same map and different reduce implementation such as count resulted in same intermediate data but was used for internally between two phases. Thus the intermediate remained unused and processing was carried out on periodically increasing raw data affecting the processing time.

B. Decision making based on pattern matching

The decision making phase gets a single document from raw data and performs map only operation and stores the result in scratch database. The resultant output's keys are compared with the key of intermediate collection.

Scratch collection: [_id, value]

Intermediate collection: [_id, value]

On matching the mapreduce operation of the workflow is performed on the intermediate collection where original map function is altered to logic given below:

```
for(var index in this.value) {

        emit (this.key ,this.value[index]);

    }
```

The reduce phase is kept same as the original one. This overall reduces the processing incurred that would be large when executed on raw data collection.

C. Intermediate Data and Result in Proposed Methodology

Experiment with proposed system at only first time the map function of map reduce is executed on raw data collection, the successive mapreduce jobs were first sent to decision making phase where Algorithm 1 is applied to the minimal number of document using query parameter and storing the resultant output in scratch collection. This result is compared with the required pattern, on success map function was altered and output was stored in intermediate collection. The collection was also altered on which the mapreduce to be performed. If the pattern did not match the normal execution was performed. In effect of this the input size decreased for same calculation, intermediate data got reused and processing time was also reduced compared to existing mapreduce workflow. Figure 9 shows the comparison of input size and processing time of existing and proposed methodology.

The result of analysis:

P=Number of documents in raw data collection.

I=Number of documents of intermediate data.

Size=Input size for map phase

1. Time(I) < Time(P)
2. Size(I) < Size(P)

TABLE I

INPUT SIZE FOR EXPERIMENT AMONG SYSTEMS.

Sr no.	Existing System		Proposed Methodology	
	Number of documents	Processing time (ms)	Number of documents	Processing time (ms)
1	500000	9960	25000	3174
2	1000000	20305	50000	5639
3	1500000	30314	75000	8511
4	2000000	39056	100000	10969

Table I. shows the number of documents and corresponding processing time in mapreduce for existing and proposed methodology. We can see that there is a drop in the number of documents to be processed which would decrease the execution time. As the number of documents to be processed decreases there is a relative decrease in processing time this is shown in Figure 9 which shows the execution time between the systems.

VI. CONCLUSION

The proposed methodology and the experiment performed results that for large datasets the intermediate data storage and re-usage of can be highly efficient for mapreduce framework. The limitations described in section II(c) are overcome using proposed mechanism. It was shown how changes in the map side affect the processing. As per general implementation large calculations are carried out on map phase, so diverting the implementation on collection having intermediate data bring down the complex processing to minimal processing leading to reuse of intermediate data and managing of input size with a reduction in overall processing time. The challenge with the proposed model is that the intermediate data produced by the mapper has some extra documents internally generated which exceed the size of the collection of intermediate dataset. To overcome this jsMode was enabled but it's limited to 5, 00,000 key of result set. This is due the current version of mongodb 2.4 that does not support for multiple valued in mapreduce.

Future work for this is to perform same implementation combining Hadoop and MongoDB. Text search on Hadoop is the next challenge as there is a lot of potential areas to be worked upon to optimize and get accuracy in search in big data.

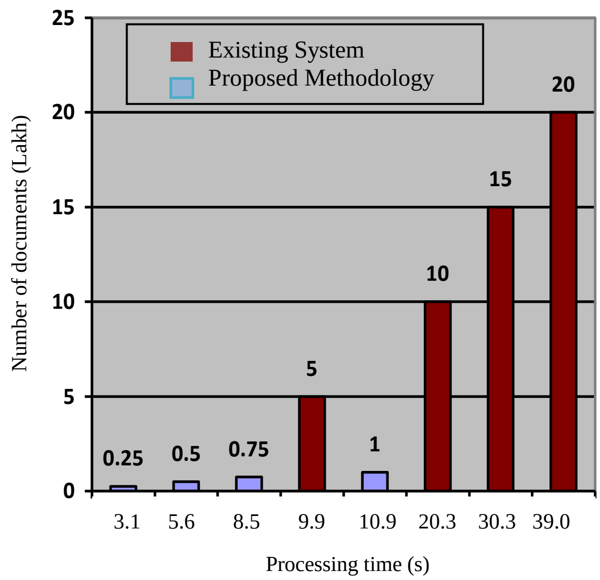


Fig 9. Number of documents (lakh) being processed and processing time (ms) in existing system and proposed methodology.

REFERENCES

[1] A. Espinosa, P. Hernandez, J.C. Moure, J. Protasio and A. Ripoll . "Analysis and improvement of map-reduce data distribution in mapping applications". Published online: 8 June 2012 JSupercomput (2012) 62:1305-1317.

[2] Diana Moise, Thi-Thu-Lan Trieu, Gabriel Antoniu and Luc Bougé. "Optimizing Intermediate Data Management in MapReduce Computations". CloudCP '11 April 10, 2011 Salzburg, Austria ACM 978-1-4503-0727-7/11/04.

[3] Iman Elghandour and Ashraf Aboulnaga. "ReStore: Reusing Results of MapReduce jobs". The 38th International Conference on Very Large Data Bases, August 27th 31st 2012, Istanbul, Turkey. Proceedings of the VLDB Endowment, Vol. 5, No. 6.

[4] J. Dean and S. Ghemawat. "MapReduce: Simplified data processing on large clusters". In Proc. OSDI, pages 137-150. 2004.

[5] Qiang Liu, Tim Todman, Wayne Luk and George A. Constantinides. "Automatic Optimisation of MapReduce Designs by Geometric Programming". FPT 2009 IEEE.

[6] Rabi Prasad Padhy. "Big Data Processing with Hadoop-MapReduce in Cloud Systems". International Journal of Cloud Computing and Services Science (IJ-CLOSER) Vol.2, No.1, February 2013, pp. 16~27 ISSN: 2089-3337.

[7] Raman Grover and Michael J. Carey. "Extending Map-Reduce for Efficient Predicate-Based Sampling". 2012 IEEE 28th International Conference on Data Engineering.

[8] Ruxandra Burtica, Eleonora Maria Mocanu, Mugurel Ionut, Andreica, and , Nicolae Ţăpuş. "Practical application and evaluation of no-SQL

databases in Cloud Computing". 2012 IEEE. This paper is under research grants PD_240/2010 (AATOMMS - contract no. 33/28.07.2010) from the PN II - RU program and ID_1679/2008 (contract no. 736/2009) from the PN II - IDEI program.

Shah Pratik Prakash , School of Computing Science and Engineering
VIT University – Chennai Campus. Chennai, India.
pratik.prakash2013@vit.ac.in
Pattabiraman. V, School of Computing Science and Engineering
VIT University - Chennai Campus. Chennai, India.
pattabiraman.v@vit.ac.in

Statistical Analysis of Different Artificial Intelligent Techniques applied to Intrusion Detection System

Hind Tribak, Olga Valenzuela, Fernando Rojas, Ignacio Rojas.
University of Granada. Spain

Abstract— Intrusion detection is the act of detecting unwanted traffic on a network or a device. Several types of Intrusion Detection Systems (IDS) technologies exist due to the variance of network configurations. Each type has advantages and disadvantage in detection, configuration, and cost. In general, the traditional IDS relies on the extensive knowledge of security experts, in particular, on their familiarity with the computer system to be protected. To reduce this dependence, various data-mining and machine learning techniques have been used in the literature. The experiments and evaluations of the proposed intrusion detection system are performed with the NSL-KDD intrusion detection dataset. We will apply different learning algorithms on NSL-KDD data set, to recognize between normal and attack connections and compare their performing in different scenarios- discretization, features selections and algorithm method for classification- using a powerful statistical analysis: ANOVA. In this study, both the accuracy of the configuration of different system and methodologies used, and also the computational time and complexity of the methodologies are analyzed.

Keywords—Intrusion detection, NSL-KDD intrusion detection dataset, learning algorithms, ANOVA.

I. Introduction

Security is an important issue for all the networks of companies and institutions at the present time and all the intrusions are trying in ways that successful access to the data network of these companies and Web services and despite the development of multiple ways to ensure that the infiltration of intrusion to the infrastructure of the network via the Internet, through the use of firewalls, encryption, etc. But IDS is a relatively new technology of the techniques intrusion detection methods that has emerged in recent in recent years. Intrusion detection system's main role in a network is to help computer systems to prepare and deal with the network attacks. An intrusion detection system is a component within the security model of an organization. Consists in detecting inappropriate, incorrect or anomalous activities from the outside-inside, of a computer system [1].

Intrusion detection system is classified into two categories: signature based detection systems and anomaly based detection systems. Signature based detection system (also called misuse based): This type of detection is very effective against known attacks, and it depends on the receiving of regular updates of patterns and will be unable to detect unknown previous threats or new releases.

Anomaly based detection system: This type of detection depends on the classification of the network to the normal and anomalous, as this classification is based on rules or heuristics rather than patterns or signatures and the implementation of this system we first need to know the normal behavior of the network. This type use learning to build the normal behavior profile of the system.

II. Data preprocessing techniques

The pre-processed data is an important phase in the data mining, which is explained in the next paragraph, thanks to which is meant to prepare the data for the phase of knowledge discovery, and to her be quick and easy. Some of the techniques used are:

A. Feature Selection

It is a data reduction technique. A preconstruction stage of the classifier is to determine which attributes or characteristics of the data set are relevant, to carry out a good differentiation between classes [2]. The process by which these attributes are selected is called feature selection. This process consisted of 4 key stages: Selection Procedure: This stage determines the possible subset of features to perform the representation of the problem. Function Assessment: This stage evaluates the subset of features selected in the previous section. Stopping criterion: it checks whether the selected subset satisfies the criteria for stopping the search. Validation Procedure: This stage is used to verify the quality of the subset of features that were identified. Depending on the way of evaluation, they are classified into 2 groups: Filter, which uses a heuristics to determine to what degree are related, the attribute and the class

This work has been partially supported by the projects: Spanish Ministry of Science and Innovation SAF2010-20558; Genil Start-Up Projects For Young Researchers PYR-2012-8 and Excellence projects Junta de Andalucia P09-TIC-5476,P07-TIC-02906.

First, Third and Fourth Authors are with the Dpt. Computer Architecture and Computer Technology. University of Granada, Spain.

Second Authors is with Dpt. Applied Mathematics. University of Granada Spain.

ISBN: 978-1-61804-227-9

that owns the data, and wrappers that make use of a classifier to evaluate the subset of attributes.

We will use two Filters method, one CFS [3] (based correlation filter) and the other the CNS. This selection of variables is based on the correlation of patterns with class variable. This method tries to find the optimal subset of attributes highly correlated with the class and at the same time, with a low degree of redundancy between them. The other method, CNS [4], look for combinations of attributes whose values divide the data into subsets containing a strong single class majority. Usually the search is biased towards small feature subsets with high class consistency. For wrapper method's we choose Naïve Bayes and a C4.5 classifier.

B. Discretization Methods

This is a data reduction technique by which it intends to build models of classification compact and simple. The discretization is very useful because thanks to it, algorithms that use qualitative or categorical variables can be used on numeric data sets on the other hand, algorithms that handle discrete data, can be evaluated on data sets with continuous variables.

There are various classifications of classification techniques, but that we will use, can be categorized into two groups: supervised and unsupervised, the difference is that the first use the class information to create breakpoints. We use two different methods. The discretization method of Fayyad and Irani [6], which works without a predefined number of intervals. During the training phase discretization a table for each of the measured characteristics is constructed. This table provides a set of numerical levels for each feature, later the actual values will be replaced by ranges to Fayyad which they belong.

This table is obtained by distributing the discretization of numeric attributes values obtained in this phase in a set of ranges. The algorithm uses the supervised discretization method based on MDL ("Minimum Description Length"), also known as entropy-based discretization [5], which is measured as where m is the total number of intervals, pi are the probabilities of different codes, and b is the unit of measure.

The feature is recursively divided into intervals in each phase with respect to minimizing the entropy of the intervals and the information required to specify these intervals. Separation is stopped when the entropy cannot be further reduced. On the other side, the other method use is the so called: Equal frequency intervals. This method is an unsupervised method which basically operates as follows: requires a feature's values to be sorted, therefore assuming that the discretized attribute has m distinct values, this method divides the domain of each variable into n parts, and each part has m/n continuous values of the attribute.

III. LEARNING ALGORITHM BASED ON ARTIFICIAL INTELLIGENT TECHNIQUES

Data mining has been defined as "The nontrivial extraction of implicit, previously unknown, and potentially useful information from data"[7]. It uses machine learning, statistical and visualization techniques to discovery and present knowledge in a form which is easily comprehensible to humans. The DM contains many study areas such as machine–learning, pattern recognition in data, databases, statistics, artificial intelligence, data acquisition for expert systems and data visualization. The most important goal here is to extract patterns from data and to bring useful knowledge into an understandable form to the human observer. Typically, a data mining algorithm constitutes some combination of the following three components:

1.- The model: The function of the model (e.g., classification, clustering) and its representational form (e.g. linear discriminants, neural networks). A model contains parameters that are to be determined from the data.

2.-The preference criterion: A basis for preference of one model or set of parameters over another, depending on the given data.

3.- The search algorithm: The specification of an algorithm for finding particular models and parameters, given the data, model(s), and a preference criterion. Though, there are lots of techniques available in the data mining, few methodologies such as Artificial Neural Networks, K nearest neighbor, K means approach, are popular currently depends on the nature of the data. Below we will present learning algorithms that have been used.

A. Genetic Algorithm

The Genetic Algorithm (GA) is a search heuristic that mimics the process of natural evolution. This heuristic is routinely used to generate useful solutions to optimization and search problems. Genetic algorithms belong to the larger class of Evolutionary Algorithm (EA), which generates solutions to optimization problems using techniques inspired by natural evolution, such as inheritance, mutation, selection and crossover.

B. Artificial Immune Systems

The biological immune system is a robust, complex, adaptive system that defends the body from foreign pathogens. It is able to categorize all cells (or molecules) within the body as self-cells or non-self cells The immune Network theory had been proposed in the mid-seventies[8]. The hypothesis was that the immune system maintains an idiotype network of interconnected B cells for antigen recognition. These cells both stimulate and suppress each other in certain ways that lead to the stabilization of the network. Two B cells are

ISBN: 978-1-61804-227-9

connected if the affinities they share exceed a certain threshold, and the strength of the connection is directly proportional to the affinity they share. The algorithm used in this paper has been CLONALG[9].

C. Artificial Neural Network

Artificial Neural Networks (ANN) is systems inspired by the research on human brain[10]. Artificial Neural Networks (ANN) networks in which each node represents a neuron and each link represents the way two neurons interact. Each neuron performs very simple tasks, while the network representing of the work of all its neurons is able to perform the more complex task. A neural network is an interconnected set of input/output units where each connection has a weight associated with it. The network learns by fine tuning the weights so as able to predict the call label of input samples during testing phase. For our experiments we have used MLP- Multilayer Perceptron-. Besides we have evaluated our data set using RBF-Radial Basis Networks-.

The idea of Radial Basis Function (RBF) Networks derives from the theory of function approximation. The Multi-Layer Perceptron (MLP) networks with a hidden layer of sigmoid units can learn to approximate functions. RBF Networks take a slightly different approach. Their main features are: a) They are two-layer feed-forward networks; b) The hidden nodes implement a set of radial basis functions (e.g. Gaussian functions); c) The output nodes implement linear summation functions as in an MLP; d) The network training is divided into two stages: first the weights from the input to hidden layer are determined, and then the weights from the hidden to output layer; e) The training/learning is very fast; f) The networks are very good at interpolation.

D. SVM

The theory of SVMs was initially developed by V. Vapnik [11] in the early 80's and focuses on what is known as Statistical Learning Theory. SVM initially appeared to separate two classes, although its application has extended to any finite number of classes. These techniques perform a linear classification, upon vectors processed on a higher dimensional space, ie in the transformed space, separating the different classes using an optimal hyperplane to maximize the margin between classes. We have evaluated SVM using different algorithms like SMO-Sequential Minimal Optimization- and using LIbSVM [12]library, both with different types of kernel function.

E. Fuzzy Logic

Fuzzy control systems, can describe the set of rules, which would use a human being, who control the process, with all the inaccuracies that have languages natural. Fuzzy logic relaxes the idea of membership of an element to a set. In traditional logic, an element belongs or not to a set, however, fuzzy logic, an element belongs to a set with a certain degree of membership. In the detection of intruders, fuzzy logic can be applied in various ways. One of them is in learning systems with fuzzy rules. This is the case of algorithm Furia[13].

F. Decision Trees

A classification tree consists of nodes, arcs and leaves. Each node represents a decision on a particular attribute value, being the terminal nodes where a decision is made about the class map. When classifying a new case will have to be compared the values of the attributes with the decisions taken at the nodes, following the branch, that match those values, in every decision. Eventually you will reach a terminal node or leaf that predicts the class for the case treated. One of the algorithms based on decision trees most used is the C4.5 [14] whose approach is TDIDT (Top Down Induction of Decision Trees), which is characterized by using a strategy of divide and conquer descending. Another algorithm, is the Random Forest introduced by Breiman in 1999, uses a set of trees (forest) classification. To classify a new data, is taken as input for each tree and produces the corresponding output classification. As a final decision for the whole of trees, takes the class with the most votes [15]. Other algorithms have been evaluated like CART[16]- Classification and Regression Trees- and NBTREE[17] which is a Naïve Bayes decision tree.

G. Rule Induction

Rule induction algorithms offer an approach to data-driven knowledge discovery from labeled data. Patter classifiers that are induced by rule learning algorithms are often simpler and easier to comprehend by humans than those induced using neural network or Support vector machines. We have used two algorithms for rule induction like RIPPER [18], or PART[19]- which build partial decision trees to induce a rule- to evaluate its performance over the data set NSL.

H. Nearest Neighbor

The basics of neighborhood classification were established by [20] in the early 50's. The nearest neighbor method and its variants are based on the intuitive idea that similar objects belong to the same class, so that the class to which it belongs an object can be inferred from the class they belong objects of the learning sample that most resembles him. The idea of similarity is reflected formally in the concept of distance. The algorithm k-NN [21], is included within the so called lazy learning techniques, and does not generate a knowledge structure, which shapes the information inherent in the training set, but the dataset itself, represents the model, i.e. is not built any model, the model is the database itself or the training set. We have evaluated KNN algorithm using 2 different variants, one using k= 1 and the other one with k =50.

I. Bayesian Networks

A Bayesian network is a directed acyclic graph and annotated, describing the joint probability distribution that governs a set of random variables. The topology or the network structure not only provides information on probabilistic dependencies between variables, but also on the conditional independence of a variable or set of them, given one or more other variables. Provide flexible methods of reasoning based on the propagation of probabilities over the network in accordance with the laws of probability theory. The algorithm TAN (Tree Augmented Naive) [22], builds a classifier, where there is a tree structure of dependencies among the predictors. Based on a model similar to dependencies Naïve Bayes, adding conditional dependencies between nodes, the algorithm TAN forming a tree between them. The mixture of both strategies enables the relaxation of independence between the predictor variables. This model, proposed by Friedman [22], is based on the calculation of the conditional mutual information between pairs of variables. This is a probabilistic paradigm, we have selected Naïve Bayes algorithm and TAN algorithm to evaluate NSL data set.

J. Hidden Markov Models

Introduced by L. E. Baum in the 70's [23], Baum proposes this model as a statistical method of estimation of probabilistic functions of a Markov chain. A HMM can be represented as a directed graph of transitions / emissions. A discrete hidden Markov model is defined in terms of the following elements [24]: 1. Number of states model "N" 2. Number of observations different "M". 3 transition probability matrix A = {aij}. 4 observation probability matrix B = {bi (k)}. 5 Probability initial state $\Pi = \{\pi\}$. Because the number of hidden states can affect the performance of classification, we evaluate HMM with different numbers of hidden states for classification. The number of states taken is based on the number of attributes that different feature selection methods have resulted.

IV. DATA SET

NSL data set is based on KDD-CUP'99 data set consists of single connection vectors where each single connection vectors consists of 41 features [25]. These features had all forms of continuous and symbolic with extensively varying ranges falling in four categories:

1. In a connection, the first category consists of the intrinsic features which comprises of the fundamental features of each individual TCP connections. Some of the features for each individual TCP connections are duration of the connection, the type of the protocol (TCP, UDP, etc.) and network service (http, telnet, etc.).

2. The content features suggested by domain knowledge are used to assess the payload of the original TCP packets.

3. Within a connection, the same host features observe the recognized connections that have the same destination host as present connection in past two seconds and the statistics related to the protocol behavior, service, etc are estimated.

4. The similar same service features scrutinize the connections that have the same service as the current connection in past two seconds.

A variety of attacks incorporated in the dataset fall into following four major categories:

A) Denial Of Service Attacks (DOS): A denial of service attack is an attack where the attacker constructs some computing or memory resource fully occupied or unavailable to manage legitimate requirements, or reject legitimate users right to use a machine.

B) User to Root Attacks (U2R): exploits are a category of exploits where the attacker initiate by accessing a normal user account on the system (possibly achieved by tracking down the passwords, a dictionary attack, or social engineering) and take advantage of some susceptibility to achieve root access to the system.

C) Remote to User Attacks (R2L): takes place when an attacker who has the capability to send packets to a machine over a network but does not have an account on that machine, makes use of some vulnerability to achieve local access as a user of that machine.

D) Probes: Probing is a category of attacks where an attacker examines a network to collect information or discover well-known vulnerabilities. These networks investigations are reasonably valuable for an attacker who is staging an attack in future. An attacker who has a record, of which machines and services are accessible on a given network, can make use of this information to look for fragile points.

V. EXPERIMENTAL STUDY

As shown in Fig.1 we can see the distribution and frequency of attacks in the data set being observed that there are certain types of attacks that take precedence over other.

For this case, the attacks are mapped to the category to which they relate and the system is trained with records like "Normal", "Two", "Probe", "R2L" or "U2R" (Fig.2). This classification is the most studied in the literature.

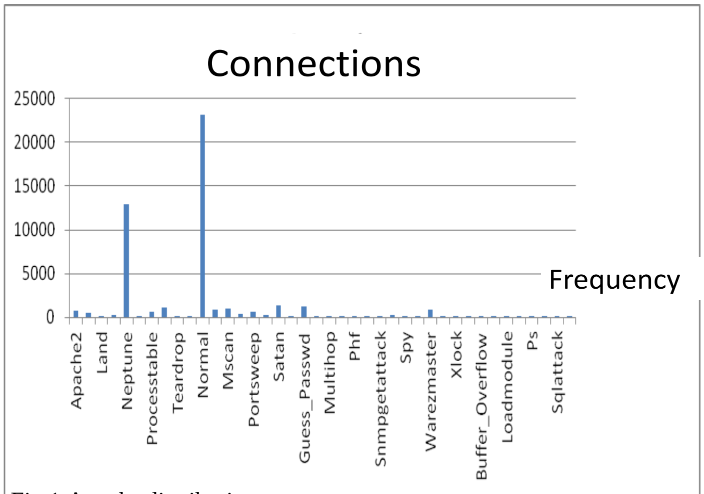


Fig.1 Attacks distribution

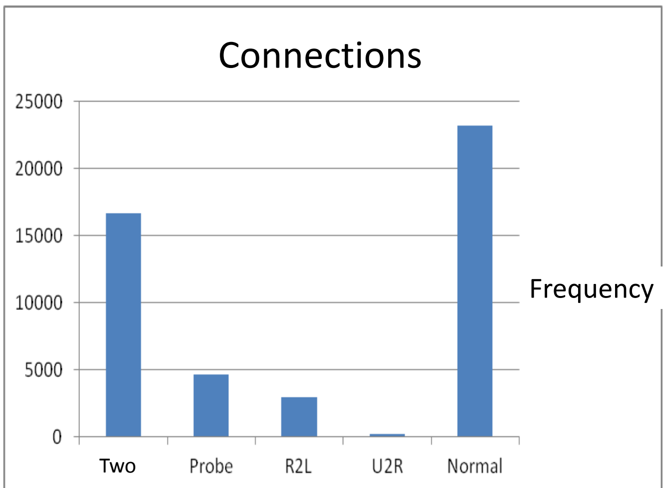
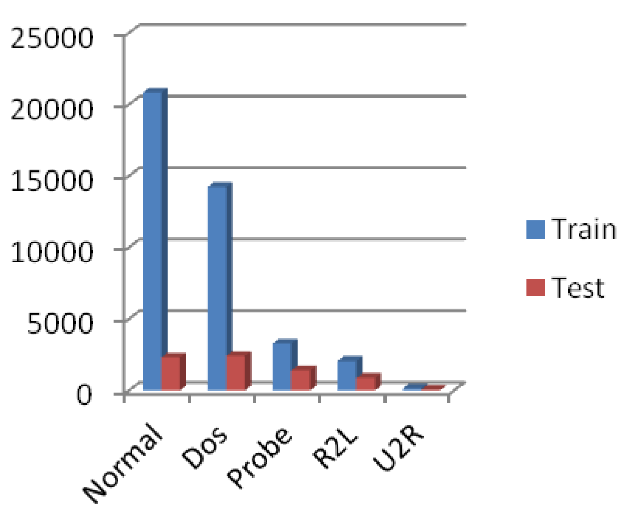
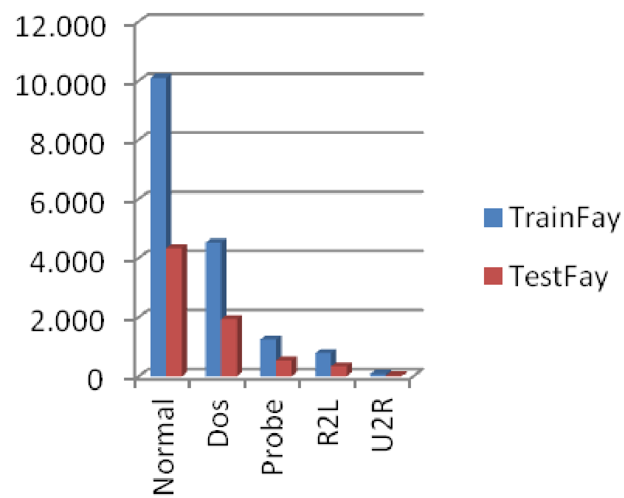


Fig.2 Attacks Mapped

As shown in the figures below we built data set without discretization, data set with Fayyad & Irani discretization and data set with Equal Frequency Intervals discretization. It is important to note that discretization significantly reduces the size of the dataset.





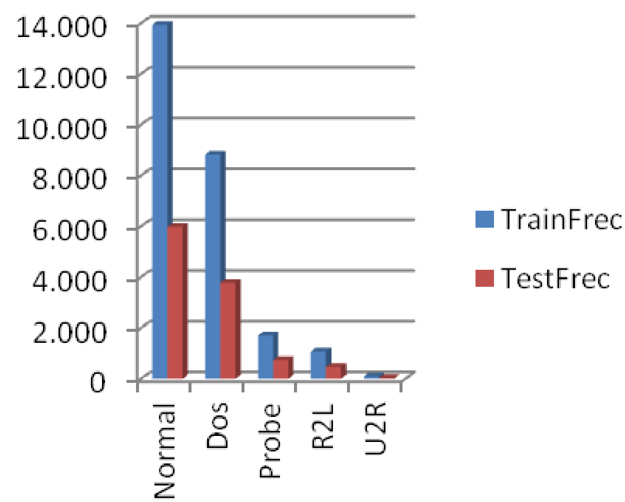


Fig.3 a) Without Discretization; b) Fayyad & Irani Discretization; c) Equal Intervals Frequency Discretization

We train and test a first kind of file without selection features and discretization, a second pair, using Fayyad Discretization and a third pair using Equal Intervals Frequency. On the other hand for each training and testing data set mentioned we will apply selection features.

The total files numbers is 30 (15 files for training and 15 for testing). For each one we will build a model using 20 different learning techniques. Finally we build a table which contains the information about the global accuracy of a classifier and the time consuming to do build it.

The evaluation of a pattern recognition experiment is based on the measure of success (percentage of well classified samples or instances) of a data set, also called the test set. Therefore in this study were obtained construction time of the models and the confusion matrices. It should be noted that 315 executions has been carried out.

ISBN: 978-1-61804-227-9

VI. Statistical Analysis

We will study how the selection of different values or levels for:

a) Filter Type;
b) Discretization Type;
c) Algorithm Type

Influence on two output variables:

a) Error classification system global (denoted as AcGlobal)
b) Time required to build the model (denoted as TimeTr).

Filter, discretization and algorithm are called a factor. A factor is an independent treatment variable whose settings (values) are controlled and varied in the experiment. The different values settings are called levels.

For Filter are: 1) All: all the features (no filter is applied); 2)FC4.5: filter C4.5 (25 features are used); 3)FCFS: filter CFS is used (10 features); 4)FCNS: filter CNS (15 features); 5) FNB: filter Naive Bayes is used (11 features).

For Discretization are: 1) 0: means that no discretization is performed; 2) Fay: discretization of Fayyad and Irani; 3) Fre: Equal frequency intervals is used.

Finally, for the type of intelligent classifier or algorithm, the levels are the following 20 methods: PML-multilayer Pervceptron-, RBFNet (Radial basis Network), Bayesian algorithm (Naïve Bayes and TAN), FURIA (fuzzy), Rule induction (RIPPER and PART), K-nearest neighbor (KNN with 1 and 50 neighbors), Markov-11, Markov-15, Markov-25, (Hidden Markov Model with different parameters), Decision Trees (Random Forest, SimpleCART, NBTree and C4.5), Genetic Algorithm, Clonalg(Immune) , and SVM (SMO using polynomial (SMOPoly) and Gaussian (SMORBF) kernel, and Libsvm with sigmoid kernel (C-SVCSigmoide) and gaussian (C-SVCRBF)).

To know if a factor has an influence or not on the output variable we analyze ANOVA table. This procedure execute the ANalysis Of VAriance of the different factors for ACGlobal and TimeTr.

Main effects	*Sum of squares*	*D.F.*	*Mean square*	*F-ratio*	*P-value*
A:Filtro	827,943	4	206,986	2,46	0,0458
B:Discr	5438,28	2	2719,14	32,29	0,0000
C:ALG	41924,9	21	1996,42	23,71	0,0000
Residue	24166,6	287	84,2041		
TOTAL	71471,6	314			

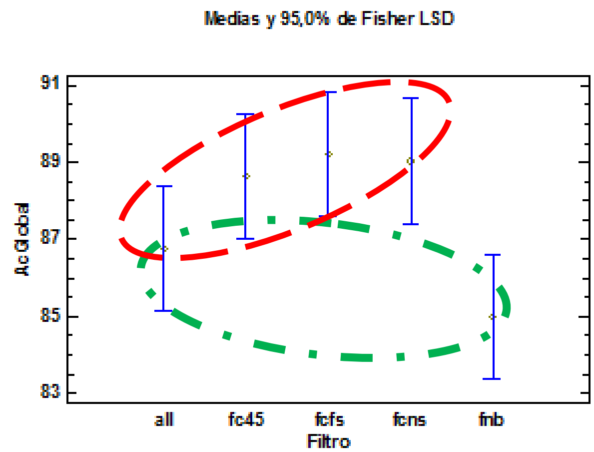
Table 1. ANOVA table analysis for Acglobal

Main effects	*Sum of squares*	*D.F.*	*Mean square*	*F-ratio*	*P-value*
A:Filtro	2,23405E8	4	5,58512E7	1,33	0,2586
B:Discr	1,92204E8	2	9,61022E7	2,29	0,1031
C:ALG	5,59739E9	21	2,66543E8	6,35	0,0000
Residue	1,20455E10	287	4,19703E7		
TOTAL	1,80374E10	314			

Table 2. ANOVA table analysis for

For both cases, we can see that Filter, Discretization and Algorithm have the value P (in red) more less than 0.05, that mean this factors have an statistically significant effect on AcGlobal and TimeTr (output variables) with a 95% confidence level.

The following figure (Fig.4) shows how the variables Filter and Discretization influence on AcGlobal.



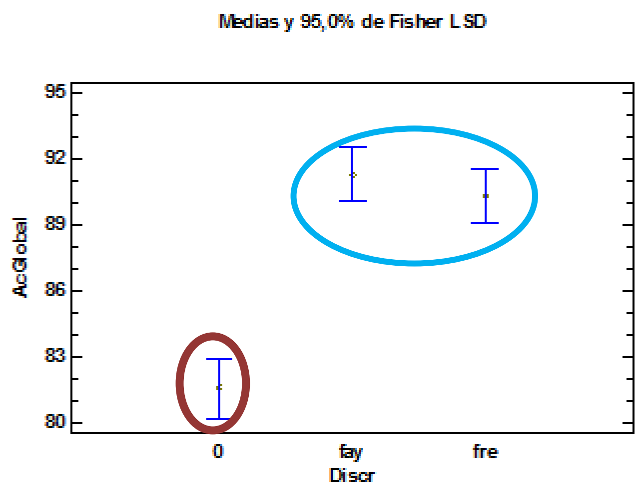


Fig.4 a) Evolution of the AcGlobal for different level of factor Filter; b) Evolution of the AcGlobal for different level of factor Discretization.

The discontinuous lines represent groups of levels that are homogeneous (i.e. this levels from a statistical point of view have the same repercussion on the output variable).

We can see that there are two homogeneous groups for the filter factor with intersection, one with FNB filter and ALL, and the other with the rest of filters and ALL. From the statistical point of view, the worst performance is for the first group which include FNB (wrapper with Naïve bayes) and ALL (without selection features). The other filters CFS (correlation filter), CNS(consistency filters), FC4.5 (wrapper with C4.5) and All (without filter) for the output variable, overall accuracy, these four types of filters are equivalent or similar, and have the same behavior.

For the discretization methods, the worst behavior is for 0 – without discretization- and Fayayd & Irani and equal frequency interval have the best. The different between the use of discretization or not, is significant.

Finally, for the algorithm used for the classification, there are 10 different groups:

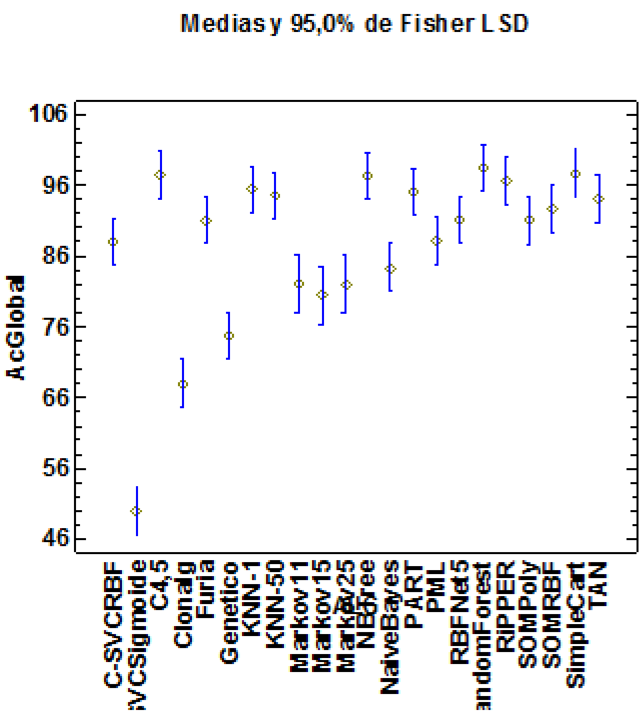


Fig.5 Evolution of the AcGlobal for different level of factor intelligent classifier.

As can be seen in Fig.5, there are ten different groups in which you can group the algorithms used, mentioning that there are no statistically significant differences between those levels sharing the same column of X's. As some methods belong to more than one group, therefore exist intersection between them.

In the first group, comments that the method: C-SVCSigmoide, SVM with sigmoid kernel. In the second group: Artificial Immune Algorithm Clonalg-. In the third group the different variants of Markov models: MARKOV15, MaARKOV25 and MARKOV11, along with the genetic algorithm.

Successively we can mention the other groups until the last that would have the algorithms: SOMRBF, TAN, KNN-50, PART, KNN-1, RIPPER, NBTREE, C4.5, CART, RANDOM FOREST, this group is the best results obtained. As important conclusion, the best algorithm is RANDOM FOREST.

VII. Conclusion

We have a set of 15 different files with a combination of selection features methods, discretization and without them. On the other hand we have carried out 315 different executions with different algorithms and in some cases its variants.

And our goal has been study how this factors influence on the accuracy of classifier and the time to build it.

We have many techniques to help us to build a good model to classify different kind of attacks into 5 different categories.

Thanks to ANOVA we have demonstrated that The discretization method (supervised, unsupervised and without discretization), the selection of attributes method (filter, wrapper, and without using feature selection) and the selected algorithm, influence the behavior of the classifier, both, from the point of view of the error and the time required to obtain the build the model.

For example, for the error of the classifier, ANOVA revealed that, the unsupervised discretization type -FRE,-equal frequency intervals-, and the supervised Fayyad method improves the accuracy of the system. Regarding the features selection, wrapper method based on Naive Bayes got worse performance and the CFS- filter method- has improved it. It was found that the filter methods have had better results. In this first case we can say that the selection features provides significant improvement in the classification task. Finally regarding the algorithm employed, have been shown that the simple methods like, Decission trees and nearest neighbor have shown better results than other more complex methods. Random Forest algorithm has had better performance.

The same methodology was used for the computational time, observing that in this case the most significant conclusion was extracted from the method used to build the classifier. In this case, MLP- Multilayer perceptron- is the most time-consuming.

The most important conclusion here is that we have used different strategy to build robust and reliable systems to detect intrusion, attacks or threats, and the result have shown that simple and straightforward techniques have obtained the best results, like Decision trees (Random Forest, C4.5) followed by induction of rules and KNN methods.

References

[1] Jian Pei, Shambhu J. Upadhyaya. Faisal Farooq, Venugopal Govindaraju “Data mining for intrusion detection: techniques applications and systems “ In Proceedings of the 20th International Conference on Data Engineering, pp: 877 - 87, 2004

[2] Dash, M., Liu, H., 1997. Feature selection for classification. Int. J. Intell. Data Anal., 1:131-156.

[3] Hall, M. A. & Smith, L. A. Feature Selection for Machine Learning: Comparing a Correlation-Based Filter Approach to the Wrapper in `Proceedings of the Twelfth International Florida Artificial Intelligence Research Society Conference, Orlando, USA' pp. 235-239. 1999

[4] H. Liu and R. Setiono. A probabilistic approach to feature selection: A filter solution. In Proceedings of the 13th International Conference on Machine Learning, pp 319–327. Morgan Kaufmann, 1996.

[5] U.M. Fayyad and K.B. Irani. Multi-interval discretization of continuous valued attributes for classification learning. In Proceedings of the 13th International Joint Conference on Artificial Intelligence, pp 1022-1027, Morgan Kaufmann, 1993.

[6] Ellis J. Clarke, Bruce A. Barton,"Entropy and MDL Discretization of Continuous Variables for Bayesian Belief Networks" International Journal Of Intelligent Systems, VOL. 15, 61]92 _2000.
[7] W. Frawley and G. Piatetsky-Shapiro and C. Matheus, Knowledge Discovery in Databases: An Overview. AI Magazine, Fall 1992, pgs 213-228.
[8] Jerne, N. K. (1974). Towards a Network Theory of the Immune System, Ann. Immunol. (Inst. Pasteur) 125C, 373 – 389
[9] L. Ñ unes de Castro and J. Timmis. An artificial immune network for multimodal function optimization. In Proccedings of the 2002 Congress on Evolutionary Computation (CEC'2002), volume 1, pp 669–674, Honolulu, Hawaii, May 2002.
[10] Hammerstrom, D., 1993. Neural networks at work, IEEE Spectrum, June, 26–32
[11] Vapnik, V. N. The Nature of Statistical Learning Theory. Springer. 1995
[12] Software available at http://www.csie.ntu.edu.tw/_cjlin/libsvm.
[13] Jens Christian Hühn and Eyke Hüllermeier. Furia: an algorithm for unordered fuzzy rule induction. Data Mining and Knowledge Discovery, 19(3):293–319, 2009.
[14] Quinlan J. C4.5: Programs for machine learning, Morgan Kaufmann Pub., 1993 (ISBN: 1558602380).
[15] Leo Breiman. Random forests. Machine Learning Journal, 45:5–32, 2001
[16] L. Breiman, J. Friedman, R. Olshen, and C. Stone. Classification and regresion trees. Wadsworth Int. Group, Belmont, CA, 1984.
[17] R. Kohavi. Scaling up the accuracy of naive-Bayes classifiers: a decision-tree hybrid. In Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining, pp 202-207, 1996.
[18] William W. Cohen. Fast effective rule induction. In In Proceedings of the Twelfth International Conference on Machine Learning, pp 115–123. Morgan Kaufmann, 1995.
[19] Frank, E. Y Witten, I.H. (1998): "Generating Accurate Rule Sets Without Global Optimization", en J. SHAVLIK (ed.): Proceedings of the Fifteenth International Conference on Machine Learning, Madison, Wisconsin. Morgan Kaufmann, San Francisco, pp. 144-151
[20] E. Fix and J. Hodges. Discriminatory analysis, nonparametric discrimination consistency properties. Technical Report 4, US Air Force, School of Aviation Medicine, Randolph Field, TX, 1951
[21] T. M. Cover and P. E. Hart. Nearest neighbor pattern classification. IEEE Transactions on Information Theory, IT-13(1):21–27, 1967
[22] Friedman, J. S., C. A. Tepley, P. A. Castleberg, and H. Roe, Middle-atmospheric Doppler lidar using an iodine-vapor edge filter, Optics Letters, 22, 1,648-1,650, 1997.
[23] Baum, L. E. An inequality and associated maximizaton technique in statistical estimation for probabilistic functions of markov processes. Inequalities, 3:pp 1-8, 1972.
[24] Rabiner, Lawrence R. A tutorial on hidden markov models and selected applications in speech recognition. pp 267-296, 1990.
[25] NSL-KDD data base. http://www.iscx.ca/NSL-KDD/

A Novel Method for Path Planning of Mobile Robots via Fuzzy Logic and ant Colony Algorithem in Complex Daynamic Environments

A. Fatemeh Khosravi purian and B. Ehsan sadeghian

***Abstract*— Researches on mobile robot path planning with meta-heuristic methods to improve classical approaches have grown dramatically in the recent 35 years. Because routing is one of the NP-hard problems, an ant colony algorithm that is a meta-heuristic method has had no table success in this area. In this paper, a new approach for solving mobile robot navigation in dynamic environments, based on the heuristic feature of an optimized ant colony algorithm is proposed. Decision-making influenced by the distances between the origin and destination points and the angle variance to the nearest obstacles. Ideal paths are selected by the fuzzy logic. The proposed ant colony algorithm will optimize the fuzzy rules' parameters that have been using to On-line (instant) path planning in dynamic environments. This paper presents a new method that can plan local routs all over the area and to guide the moving robot toward the final track. Using this algorithm, mobile robots can move along the ideal path to the target based on the optimal fuzzy control systems in different environments, especially in dynamic and unknown environments.**

Keywords- Ants colony algorithm, fuzzy logic, path planning, mobile robot.

I. INTRODUCTION

Robots are essential elements in today's society. The term robot is used for a wide range of mechanical machines, which have mobility [1]. In the case of mobile robots, path planning is a major challenge. Path Planning for a mobile robot which is in an environment with various obstacles, is finding a path without obstacles from the starting point to the destination. In this regard Issue such as shortness and Simplicity of route are important criteria affecting on the optimality of selected routes [2].Today finding suitable algorithms for mobile robot navigation is one of the hot scientific topics. Mobile robots have been used extensively in various fields such as space research, nuclear industries and mines. In these applications, finding a safe route for the robot is a prerequisite for success. So finding a suitable routing algorithm for the robot to move from the starting point toward the target, without collision with obstacles in the environment is an important issue in robotics. By considering the length of the path traveled by the robot, and energy consumption and its performance time, we need to follow an algorithm that finds the shortest possible route [3], [4].

Fatemeh Khosravi Purian, Young Researchers and Elite Club, Central Tehran Branch, Islamic Azad University, Tehran, Iran, (corresponding author to provide phone: 00989131867977, E-mail: Faith.Khp7@gmail.com).
Ehsan Sadeghian, Young Researchers and Elite Club, Majlesi Branch, Islamic Azad University, Majlesi, Iran, (E-mail: Ehsan.Sadeghian@yahoo.com).

After 30 years of research in this field, we can classify Investigations into classic and heuristic researches. Classical methods for over 20 years in the field of change or combination are including: possible road map [5], map [6], and cell decomposition [7] which are based on mathematical programming. The second methods are heuristic. When the classical method of solving routing problems that have Np-hard nature, For robot motion planning encountered different difficulties, Heuristic methods inspired by nature were introduced [8]. In the exploratory classification, potential field method [9], and soft computing are including: Neural Networks, Genetic algorithm [8], [10], the Simulated Annealing algorithm [3], the Ant Colony Optimization algorithm [11], Particle Swarm Optimizer algorithm [3], [8] and fuzzy logic [11]. According to [12], [13] using Meta heuristics algorithm for routing in mobile robots is increasing. To achieve optimal path recognized environments the best option is to use evolutionary algorithms [14], [13]. In this regard, the usage of the Ant Colony Algorithm due to its discrete nature and also its relationship for efficient routing issues is very significant [12], [15]. Ant Colony Algorithm for routing needs a detailed and preset map so it cannot be used in unknown environments. Another problem in the use of the evolutionary algorithm for mobile robot navigation is that the optimization process should be done before the robot moves (off-line), and then the robot will move on the path [14]. Thus, using this algorithm (in spite of good ability in the route optimization) is not possible in many practical applications. So for mobile robot navigation in unknown environments with fixed and moving obstacles, without knowing the position and movement of the obstacles, On-line (instant) routing methods must be used. Fuzzy logic has a very good performance in an on-line node to node path planning for robots [14], [16]. In this paper, to determine the optimal path of mobile robots in an unknown environment, a new method based on determination of fuzzy rule table, which has a very important impact on a fuzzy logic system performance, have been provided. Because the routing issue is a kind of Np-hard problem, here ACO algorithm is used to determine the optimum fuzzy rule table. The main objective of this paper is to provide a new method that can plan local paths around the environment routes and guide the moving robot toward the final track. This goal is achieved by optimization of fuzzy rules table for the mobile robot motion by utilizing the ant colony algorithms. In the rest of the paper: The second section presents the proposed fuzzy-colony algorithm, and then robot's work place environments, which are fairly simple to highly sophisticated will design. In the third section, the simulation results in terms of path length and the time spent in different environments with varying complexities are presented. Furthermore, in this part examples of considered parameters for the ant colony algorithm in different environments are presented. The proposed algorithm has been discussed in Section IV. The fifth section discusses about the efficiency and effectiveness of the proposed algorithm.

II. METHOD

In [15], Routing of mobile robots has been considered in known dynamic environments in three workspaces, including relatively simple, moderately complex, and complex. Considering strengths and weaknesses of evolutionary algorithms, to improve the optimality criteria for dynamic unknown environment, three workspaces listed in [15] were used. In the proposed method, the Ant Colony Algorithm will determine the optimized elements of the fuzzy rule table. For this purpose, a number of initial and final points have been considered in different workspaces. In each iteration, each ant based on the probability relationship in the ant colony algorithm will determine an output for each of fuzzy rules and finally, a solution is produced. According to the fuzzy control system, based on the produced schedule for each ant, for all pairs of initial and final considered points, routing takes place. Finally, criteria for evaluating the solution generated by the ants are the average length of the path traveled by the moving robot for each pair of points. Overall process of the proposed method is presented as follows:

A. producing Appropriate solutions by ants

In proposed method, each ant for producing a solution for each fuzzy rule considers a membership function according to selection probability law in the ant colony algorithm, shown in Fig. 1, complete their route and finally, a solution is achieved. This choice is done according to (1), the selection probability equation.

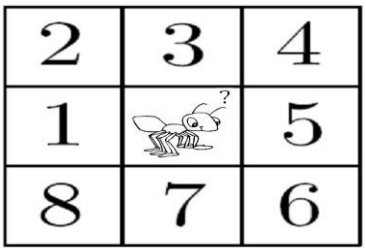


Fig. 1 Order of adjacent numbered squares

$$P_{ij}^{k}\left(t\right)=\frac{\tau_{ij}^{\alpha}\left(t\right)\times \eta_{ij}^{\beta}}{\sum_{i=1}^{n}\tau_{ij}^{\alpha}\left(t\right)\times \eta_{ij}^{\beta}} \tag{1}$$

Where τ_{ij} is the amount of pheromone in membership function for the i^{th} output for j^{th} rule. η_{ij} is heuristics data of i^{th} membership function for j^{th} rule. P_{ij} (t) is the selection probability of i^{th} membership function for j^{th} rule in iteration t. α and β are two constant parameters that indicate the amount of influence of pheromone and heuristics data in selecting the output membership functions for each rule. In This paper to speed up the convergence of the algorithm, according to (2) the table is set manually in [16] is considered as heuristics data in selection of ants. Using this relationship, selection probability of membership functions in the rules table will increase in order of 2^{β}which results in speeding up the convergence of the algorithm.

$$n_{ij}=\begin{cases}2 & if\ j^{th} MF\ G \\ 1 & Otherwise\end{cases} \tag{2}$$

B. Evaluation of the produced solutions

In each iteration, after all ants have developed their own solutions, they are evaluated according to (3), and the best ant will select.

$$Cost_{k}^{t}=\frac{\sum_{i=1}^{N}L_{i}^{k}}{N} \tag{3}$$

Where, N is the number of paths considered in different workspaces, and L_i^k is the path length for the i^{th} workspace (i^{th} initial and final point's pair) which is obtained by routing phase using the rules table for the k^{th} ant.

C. Pheromone update for all routes

After all ants have developed their own solutions and their eligibility was evaluated, general pheromone update is done according to (4). Pheromone matrix considered in this issue is a two-dimensional matrix with n rows, which are the number of output membership functions and m columns that is the number of fuzzy rules. The pheromone change for each element is updated according to (5).

$$\tau_{ij}\left(t+1\right)=\tau_{ij}\left(t\right)+\Delta\tau_{ij} \tag{4}$$

$$\Delta\tau_{ij}=-\lambda\tau_{ij}\left(t\right)+\begin{cases}\dfrac{Q}{num} & for\ nodes\ of\ the\ best\ ant \\ 0 & otherwise\end{cases} \tag{5}$$

Where $\tau_{ij}\left(t\right)$ is pheromones corresponding to i^{th} membership function for j^{th} fuzzy rule in previous iteration and $\tau_{ij}\left(t+1\right)$ is pheromones corresponding to i^{th} membership function for j^{th} fuzzy rule in current iteration. In addition, Q is the amount of pheromone secreted, and "num" is the number of ant priority among pheromones secreted candidate. λ is the pheromone evaporation coefficient which is a number between zero and one. Fig. 2 presents general flowchart of the proposed algorithm.

ISBN: 978-1-61804-227-9

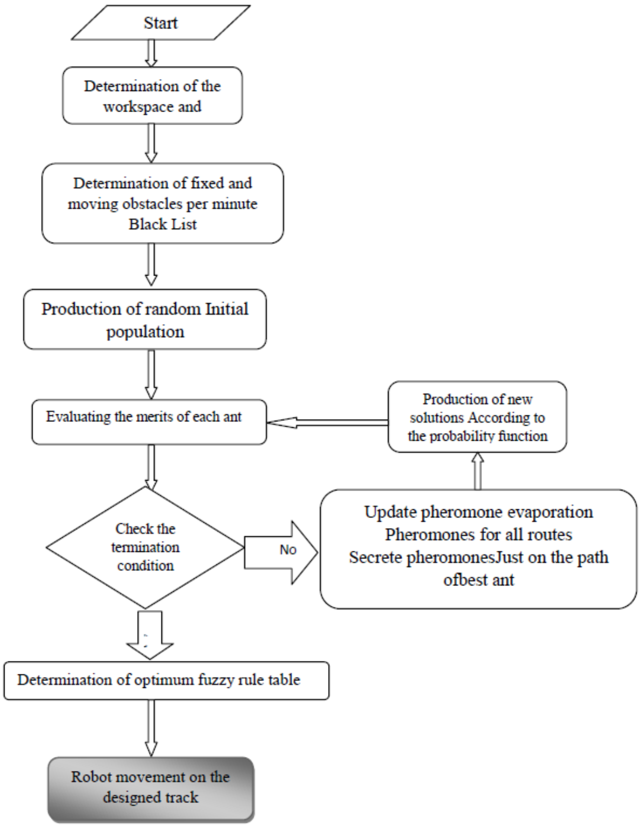


Fig. 2 Overall flowchart of the proposed algorithm.

III. THE SIMULATION RESULTS

Production of random population means each ant considers one output for each fuzzy rule then evaluation merits of each ant by considering the fuzzy rules table obtained by each ant, results in optimal routes for sets of initial and final points. The average path length for different paths is considered as a criterion for suitability evaluation of ants. It should be noted that in proposed method based on the combination of fuzzy and ant colony algorithms, routing is done only once in off-line mode, which results in optimization of parameters in fuzzy rules table. Run-time using a laptop computer with Intel Dual core processor with a processing speed of 3GHz and 4GB RAM, with Windows 7, by Matlab R2011 software, was 417 minutes, But then for routing for each specific input and output modes, Based on fuzzy logic routing is performed using the optimized rules table. In fact, with this method, the ant colony algorithm will determine the optimal elements of fuzzy rule table, with the aim to reduce the robot's path. In this paper for the simulation, three workspaces (different) with four different states (four initial and final points) are assumed. As a result, for the solution of each ant, 12 different paths are achieved by using fuzzy logic. Average length of the path found by each ant in 12 different path is considered as a criterion for suitability evaluation of each ant. Notably, then, routing based on fuzzy logic for each point in the workspace and initial and final desired points is done by utilizing the optimized rules table. As a result, the time required to implement the proposed routing algorithm based on the combination ant colony, and fuzzy algorithms is very similar to routing run-time with fuzzy logic method (manually setting the table). However, here the length of the path traveled by the robot is much shorter than the path traveled with fuzzy logic [16]. The optimal table by ant colony is according to Table1.

Table1. Table of optimized fuzzy rules obtained by Ant Colony Algorithm.

Output: priority of the next node selection	**Input 2: Angle difference with respect to target**	**Input 1: Distance to the nearest obstacle**
Very low	Very low	Very low
Very low	low	Very low
Very low	Medium	Very low
Very low	High	Very low
Very low	Very high	Very low
Medium	Very low	low
Very low	low	low
low	Medium	low
low	High	low
Very low	Very high	low
low	Very low	Medium
Medium	low	Medium
low	Medium	Medium
low	High	Medium
Very low	Very high	Medium
Very high	Very low	High
High	low	High
Medium	Medium	High
low	High	High
Very low	Very high	High
Very high	Very low	Very high
Very high	Low	Very high
High	Medium	Very high
Medium	High	Very high
Low	Very high	Very high

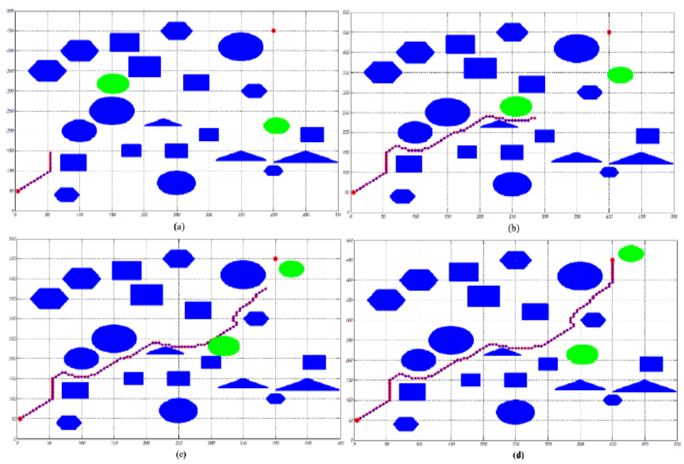


Fig. 3 Routing with the proposed method in relatively complex environments start at (50, 5) and end at (450,400), 916.24 cm path length.

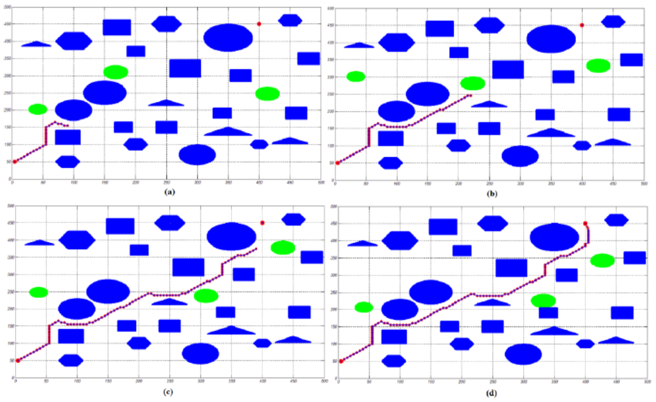


Fig. 4 Routing with the proposed method in quite complicated environment start at (50, 5) and end at (450,400), 927.18 cm path length.

In quite complicated and relatively complex workspaces in the mentioned simulated environment, the proposed algorithm improved fuzzy rules table in [16] with the evolutionary characteristics of the ant colony algorithm so that the walking robot can select the optimal path in any situation without any collision with obstacles, which may be at any speed, direction, and shape in front of robots. Table2 compares the performance of the proposed fuzzy-colony algorithm with the proposed fuzzy algorithm in [16]. In different environments path length found by the proposed fuzzy - Colony algorithm is far less than the optimal fuzzy algorithm proposed in [16]. This is due to optimization of fuzzy rules table in [16] by a sophisticated expert with evolutionary ACO algorithm. The path length is expressed in cm.

Table 2. Performance of the proposed algorithm in different environments.

workspaces complexity	Starting Point	Target Point	Fuzzy	ACO_Fuzzy optimization
Relatively simple workspaces	(50,5)	(450,400)	620.9	616.19
Relatively simple workspaces	(20,55)	(460,455)	660.52	658.69
Relatively simple workspaces	(30,65)	(465,460)	656.8	655.61
Relatively complicated workspaces	(50,5)	(450,400)	965.4	927.18
Relatively complicated workspaces	(20,55)	(460,455)	725.10	718.61
Relatively complicated workspaces	(30,65)	(465,460)	727.9	711.33
Quite complicated workspaces	(50,5)	(450,400)	924.4	916.24
Quite complicated workspaces	(20,55)	(460,455)	726.4	720.74
Quite complicated workspaces	(30,65)	(465,460)	709.6	706.75

Fig. 5 illustrates the comparison of the routes traveled by moving robots for Real-Time (on-line) fuzzy logic and optimized fuzzy algorithm with the ant colony. The improvement of the fuzzy rules table is quite evident by the ant colony algorithm. To avoid repetition, only the result of the third workspace which is quite complicated is offered.

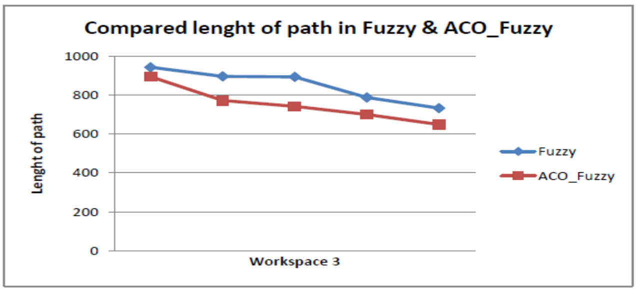


Fig. 5 Comparison of path length in quite complicated workspace by the proposed fuzzy-ant colony and fuzzy algorithms.

According to Fig. 6, it follows that for every 5 points in quite complicated environment with variable origin and target points, the ant colony offers the shortest path due to off-line search. But since all the information is not fully available to the robot, fuzzy logic approach has been proposed and by considering the length of the paths obtained, this method has the greatest distance traveled by the mobile robot. When the fuzzy rules table optimized with the evolutionary ant colony algorithm, as is also evident in figure below, The result shows that the Distance traveled by the proposed fuzzy -colony algorithm in an unknown environment is very close to the traveled distance introduced by evolutionary algorithms in [15].It should be noted that evolutionary algorithms due to prior knowledge of the environment, choose the shortest path, but in the real-world, robots are facing the unknown environment and obstacles and the proposed colony – fuzzy algorithm in such unknown environments will find short paths with respect to optimality criteria (length, distance and elapsed time). The length of these paths is close to the length of the path traveled by evolutionary algorithms that obtained from adequate and accurate information from their environment.

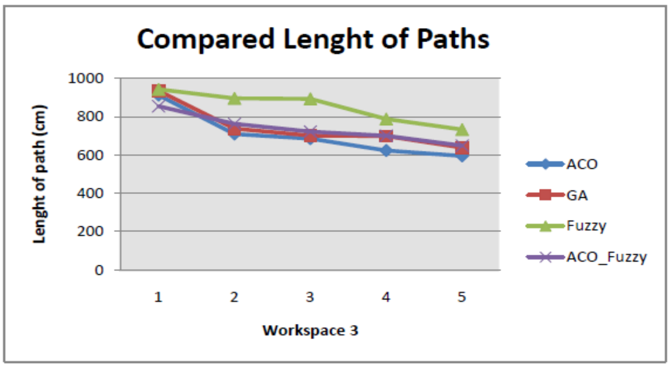


Fig. 6 Comparison of path length in quite complicated workspace by the proposed algorithms.

IV. DISCUSSION

Routing and location identifications are considered as the most important task of a mobile robot. With respect to working environment of mobile robot, different algorithms and various methods have been used. Among them for a known environment with fixed obstacles, ants routing algorithm have better performance. There are a great number of investigations in this field, but most work environments for mobile robots are areas with fixed and moving obstacles in a different direction and speed that are famous as the dynamic complicated known environments, and there is less study in this field. So in unknown dynamic environment due to lack of mapping and also shortage of mobility functions of barriers, Ant Colony Algorithm alone is not working. Because in mentioned algorithm, it is necessary to have detailed information about the environment and mobility functions of obstacles, in these cases, real-time methods must be utilized. In the meantime, fuzzy logic with fuzzy rules table plays an important role in robot's navigation in complex environments. Fuzzy rules table is usually set by an expert. On the other hand, manually setting the table cannot be fully optimized. Therefore, to optimize the parameters of fuzzy table, the evolutionary ant colony algorithm has been used in this research. In other words, the ant colony algorithm with respect to optimality criteria, affects the fuzzy table which has been set by an expert, and offers efficient routing of mobile robots. So with this method in every unknown environment a robot can find its way with the optimality criteria. In other words, in each iteration, each ant according to the selection probability relationship in ant colony algorithm, for each of the fuzzy rules, determine an output, and eventually produce a solution. Then According to fuzzy control systems based on the table produced for each ant, for all pairs of considered initial and final points routing is done.

V. CONCLUSIONS

For routing in an environment similar to the real environments on-line methods are needed. Among these methods, fuzzy logic is a good choice for solving this kind of problem, due to on-the-spot node to node routing for mobile robots and has been regarded by Researchers. Basis of fuzzy logic is determination of fuzzy rules table, which is determined by an expert. However, such a table cannot be completely optimized if it is determined manually. Therefore, considering that determining the fuzzy rules table is a kind of NP problem, using evolutionary algorithms and especially ant colony algorithm with respect to its discrete nature is very beneficial. Ant colony algorithm improves fuzzy rules table that is determined by a qualified person according to optimality criteria, which are including the length, the time and the smoothness of the path in this article and finally optimizes elements of table. As a result, the proposed hybrid algorithm enables the mobile robot to pass the optimal path at the right time to achieve the goal in every unknown environment with any obstacle.

REFERENCES

[1] I.Waheed," Trajectory / Temporal Planning of a Wheeled Mobile Robot ",Canadian Thesis in the Department of Mechanical Engineering University of Saskatchewan, December 2006.

[2] D. Karaboga et al. "Artificial bee colony (abc) Optimization Algorithm for Solving Constrained Optimization Problems", Foundations of FuzzyLlogic and Soft Computing pages 789-798, 2007.

[3] D. Xin and C. Hua-hua, "Neural network and genetic algorithm based global path planning in a static environment", Journal of Zhejiang University SCIENCE, 6A(6):549-554, 2005.

[4] Z.S. Yang. "Nature-inspired metaheuristic algorithms", Luniver press, 2011.

[5] H. DongbingGu, "A Differential Game Approach to Formation Control" *IEEE Transactions* On Control Systems Technology, Vol. 16, No. 1, January 2008.

[6] Yanyan Dai, Viet-Hong Tran, ZhiguangXu, and Suk-Gyu Lee "Leader-Follower Formation Control of Multi-robots by Using a Stable Tracking Control Method', *ICSI 2010*, Part II, LNCS 6146, pp. 291–298, 2010.

[7] SitiNurmaini ,AngginaPrimanita "Modeling of Mobile Robot System with Control Strategy Based on Type-2 Fuzzy Logic"*International Journal of Information and Communication Technology Research*, Volume 2 No. 3, March 2012.

[8] M.A. Porta Garcia, O. Montiel, O. Castillo, R. Sepu´ lveda, P. Melin, " Path planning for autonomous mobile robot navigation with ant colony optimization and fuzzy cost function evaluation", Applied soft computing Elsevier of journal, 1102-1110, 2009.

[9] P.Mellodge. "Model Abstraction in Dynamical Systems: Application to Mobile Robot Control", Thesis, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, May 10, 2007.

[10] O. Castillo and L. Trujillo, "Multiple Objective Optimization Genetic Algorithms for Path Planning in Autonomous Mobile Robots", International Journal of Computers, Systems and Signals, Vol. 6, No. 1, 2005.

[11] D. Huh, J, Park, U. Huh, H. Kim, "Path Planning and Navigation for Autonomous Mobile Robot," IECON 02 IEEE annual conference, 2010.

[12] M. Mohamad, W. Dunningan, Ant colony robot motion planning, in: The International Conference on Computer as a Tool, EUROCON, Serbia & Montenegro, pp. 213–216, 2005.

[13] O. Ob e, I. Dumitrache"AdaptiveNeuro-Fuzzy Controler With Genetic Training For Mobile Robot Control" *Int. J. of Computers, Communications & Control*, ISSN 1841-9836, E-ISSN 1841-9844 Vol. VII (2012), No. 1 (March), pp. 135-146.

[14] W. Kwong and Y. Xu, "A Generalized 3-D Path Planning Method for Robots Using Genetic Algorithm with an Adaptive Evolution Process", Proceedings of the 8th World Congress on Intelligent Control and Automation July 6-9 2010, Jinan, China.

[15] F. khosravipurian, F. farokhi, R. sabbaghi, "Comparing the performance of genetic algorithm and ant colony optimization algorithm for mobile robot path planning in the dynamic environments with different complexities", Journal of Academic and Applied Studies (JAAS) Int. Association for Academies, Canada, Vol.3, No.2, Feb 2013, ISSN: 1925-931x

[16] F. khosravipurian, E. sadeghian, "Path planning of mobile robots via fuzzy logic in unknown dynamic environments with different complexities", Journal of Basic and Applied Scientific Research, J. Basic. Appl. Sci. Res, 3(2s), 528-535, 2013.

Neural Tau (Neuτ): Hinge-pin of Neural Augmented Ant Colony Optimization (NaACO) Technique

Muhammad Umer, Riaz Ahmad, Amir Farhan

***Abstract*—**In this paper we introduce and elaborate on the concept and application of Neural Tau (neuτ) which has become the hinge pin for the formulation of Neural Augmented Ant Colony Optimization (NaACO) in our previous research and publications. The concept of neuτ has been discussed in detail and it is compared with the concept of traditional or conventional pheromone up-gradation function. The application of this entity through NaACO in various worker assignment and scheduling problems along with the areas of future research has been depicted.

***Keywords*—**Worker Assignment; Neural Augmented Ant Colony Optimization (NaACO); Artificial Neural Networks (ANN), Neural Tau (neuτ).

I. LITERATURE REVIEW

THE concept of Pheromone up-gradation in Ant ColonyOptimization (ACO) is to enable this meta- heuristic to forget the already learnt solution and enable ACO for global optima. The function of evaporation or reinforcement is thus developed and is elaborated in detail by Marco Dorigo and Thomas Stutzle [1] in their pioneering research on ant colonies and its mathematical formulation. The pheromone evaporation in real ants is not pivotal in finding the shortest path or trail for them once they are in search of food. In contrary to this once we talk about artificial ants, it is far more significant as the artificial ants have to deal with far more complex paths and problem sets as compared with the real ants. The conventional pheromone update (τ) is given as:

$$\tau_{ij} \leftarrow (1-\rho)\tau_{ij} \qquad (1)$$

Where ρ is a parameter ranging from 0 to 1. As per Dorigo and Stutzle if a single ant is taken and the "*Pheromone Depositing*" is done only on a forward or backward path, the result is that the ant colony is unable to choose the shortest branch. This impediment has to be catered for once dealing with a problem set in which inadvertently we have to use one ant to converge to the food source such as in the most common scheduling problems (problem discussed in preceding sections). As such the nature and function of the up-gradation mechanism changes and the formulation of τ has to adjust. As such τ is no longer required to upgrade but the corrective nature of τ can be used to decide upon the "*Combinational Correctness"*of the whole solution set.

Muhammad Umer is with the School of Manufacturing and Mechanical Engineering (SMME), National University of Science and Technology (NUST), Islamabad, Pakistan (email:muhammad.umar@smme.nust.edu.pk)

Riaz Ahmad is with the School of Manufacturing and Mechanical Engineering (SMME), National University of Science and Technology (NUST), Islamabad, Pakistan (email: riazcae@yahoo.com)

Amir Farhan is with Muhammad Ali Jinnah University (MAJU), Islamabad, Pakistan (email: afrafique@yahoo.com)

This capability is enforced in the conventional τ through the introduction of Artificial Neural Networks (ANNs) in the up-gradation or the corrective loop while implementing ACO. Hence we introduce **neuτ**which has become the hinge pin for our implementation of NaACO [2].

NP-complete optimization problems were the hardest or more exhaustive [3]. Branch and Bound algorithms were first used to tackle these problems. The use of operations research techniques such as dynamic programming or integer linear programming were also used to solve the basic scheduling problems. Mathirajan et al. [4] have designed heuristic methods to tackle optimization related scheduling problems.

"Intelligent scheduling" from "systems approach" can be generated in a more better and compact way through the artificial intelligence approach to scheduling problems. The basic notion of the systems approach is "Divide and Conquer", as it "conquers" the problem through amalgamation of achieved results for parts by "dividing" the problem into more realistic and manageable domains. The whole idea of implementing the AI approach to any scheduling problem is to identify the agents that are "adaptive" yet "intelligent.

The development of ACO can be associated with the theory of stigmergy by Grasse in 1959 [5]. The optimization problems were tackled by ACO through the movement of artificial ants. The most successful application of ACO was on the travelling salesman problem (TSP). In 1992, Dorigo [6] presented the first algorithm. Various other variants of ACO were introduced by Dorigo and Thomson [1]such as Ant System (AS), MIN-MAX ant systems ATNS-OAP etc. ACO has proven to be more efficient than its counterparts as shown by single machine tardiness problem [7].Likewise, flow shop optimization was also done through ACO[8]. The resource constraint problem was done by Merkle et al. [9]. A non-delay scheduling sequence was tackled by local search byBlum and Sampels[10].

Neural networks have two functions which can be used very potently to formulate a τ which is capable of learning the pattern and then participating in formulation of a Fitness Threshold for the analysis of the whole scheduling problem. The neural networks are used to memorize the information related to the problem or to remember the inputs to the problem [11]. Secondly, ANNs are also used for the optimization in a given set of constraints and objective functions.[12], [13]. The capability of self-organization, convergence through unsupervised learning, generalization of a constraint problem, fault-tolerance are further advantages not typically found in "classical" approaches. The ANN thus constructed learns the patterns based on the inputs of the scheduling problem, and the outputs. The combination of ANN with ACO for the solution of scheduling problems which are constrained by the fact that the artificial ants can only traverse in the forward direction in the first iteration, and thus this represents a hybrid approach to optimization problems. ANNs can be utilized to handle the variable of "resource" in the scheduling, in order to come up with a form of intelligent learning system or framework. Moreover, as already discussed this paradigm applies to the combinational analysis approach in scheduling problems and redefining the role of τ in the ACO approach.

II. PROBLEM FORMULATION

We have to schedule *n* jobs to m parallel machines subjects to the following conditions:

a. Each machine is capable of handling one job at a time.

b. There are three times A, B and E for each machine.

c. No job hopping is allowed.

d. Each job has to pass through a sequence of machines.

e. We have to allocate n jobs by the application of ACO and use neuτ to check for the combinational aspects of the problem through ANN.

f. The job can go to any machine first.

g. The setup times of machines are not catered for, so we are not focusing on minimizing the setup slacks.

h. Multitasking of machines is not allowed.

i. We are not accounting the transportation time taken between the machines.

j. For a start we are not focusing on adding resources in the form of machines or adding jobs to the initial problem.

k. The jobs allocation involves the calculation of the total time spent by each job on each machine.

l. At the start all machines and jobs are available.

m. The time for each machine is then calculated as;

$$P_i = A + B/E \ldots.. \quad (2)$$

Where the processing time of job *i* is given by *P* and A, B and E are respective times allocated to each machine. A set of 100 problems(figure #1) are selected each with three machines and twelve jobs. The following table represents a typical problem set:

The pseudo code for this problem is as under:

Part-I (Scheduling of Jobs)

Set n as the set of scheduled jobs for m machines. Initialize allocation of jobs through triggering heuristic.

While first loop is running apply heuristic information;

Every job has to be allocated to a machine.

Local search is to be applied.

Repeat for (n=1, 2, 3… 12);

All schedulable jobs are to be identified and partial solution is to be constructed.

Assign jobs to machines (m=1,2,3) do

end for

Part-II (Combinational Fitness)-neuτ

Train ANN for machine inputs (m=1, 2, 3) with τ values.

Display: machine times and number of workers proposed;

If YES end;

If NO do loop.

Job	Machine 1			Machine 2			Machine 3		
	A_i	B_i	E_i	A_i	B_i	E_i	A_i	B_i	E_i
1	7	556	8	8	227	1	4	328	3
2	5	784	4	7	36	5	2	330	9
3	5	195	3	3	236	9	6	570	1
4	2	427	9	9	305	4	6	260	4
5	3	85	6	8	240	7	2	506	4
6	7	799	6	0	758	5	2	166	5
7	0	540	4	9	783	5	8	148	2
8	7	12	1	3	321	5	8	466	5
9	8	460	6	5	222	4	5	64	3
10	7	80	7	7	128	4	9	366	6
11	0	82	9	0	130	3	9	724	5
12	4	639	8	5	517	1	2	209	2

Figure # 1

A. Neural Augmented ACO Formulation andResults:

Step 1

The probability by which the ants choose their path towards the food source is given by:

$$P_{i,j,k} = [\tau_{i,j}]^{\propto}[_{i,j}]^{\beta} / \sum[\tau_{i,k}]^{\alpha}[_{i,k}]^{\beta} \quad (3)$$

In ACOη is the reciprocal of the heuristic function or the greedy algorithm to trigger the initial assignment. In case of job assignment to machines the heuristic is the machine time of each machine, such that the ant (job) first goes to the machine with the shortest processing time. As a result a sequence for each job is generated for the machines (figure #2).

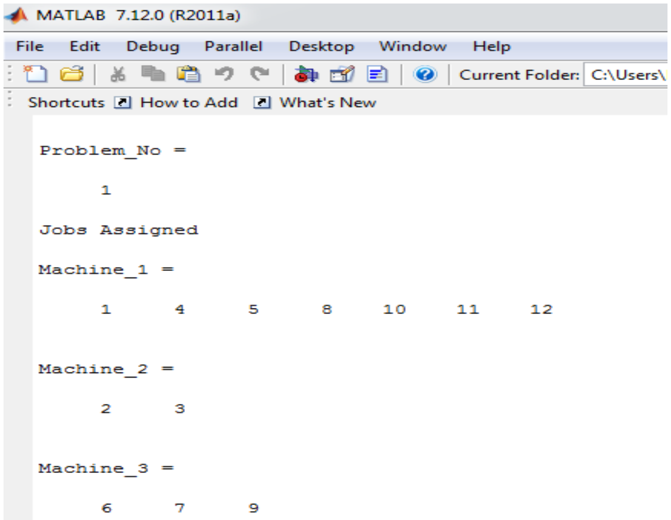


Figure #2

Step 2

A Feed Forward Neural Network is used to train the ANN in which the inputs of all 100 problems are linked with the outputs to come up with a trained ANN.Another feature is the inclusion of a YES/NO output parameter which is easy to comprehend and understand in pure industrial environments. The feed forward ANN comprises of The ANN is logic based (Yes/No) artificial neural network model, which clusters the combination of the three-machine times, initial job assigned and combines these entities to give and generate holistic optimum combination (neuτ). If the overall combination is correct the answer is "***Yes***", which means that the overall compatibility of all the variables is "recognized" and the value can be "***combinatorially used***".

Figure # 3 shows the training of ANN on the given set of 100 problems. The network is thus trained to give the NaACO approach.

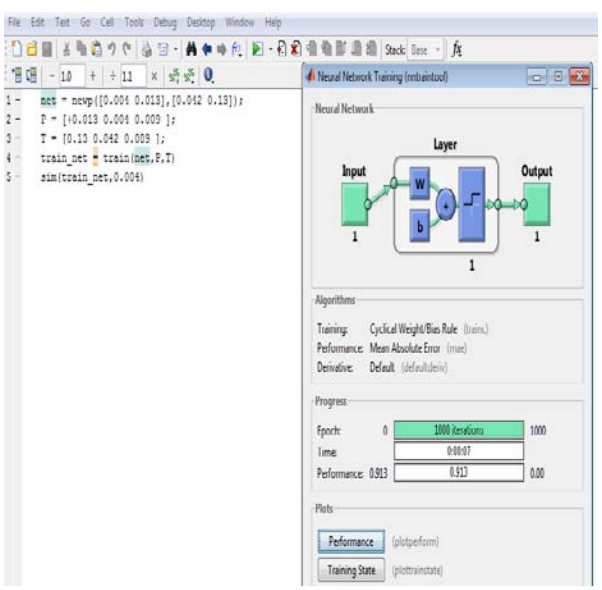


Figure # 3

This combinational optimization approach is then tested on various problem sets to check for the fitness of the combination with each other. The parameters involved include the times A, B and E on each machine, the no of jobs forecasted for that machine and the level of correctness recommended or being tested for (figure #4).

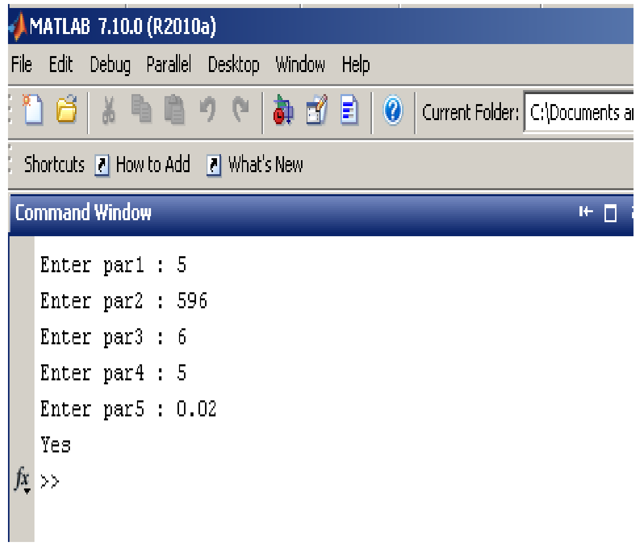


Figure #4

The overall corrective parameters of the problem set shall vary according to the complexity of the problem and according to the variables involved in the problem. ANN has to be trained to ascertain the tolerance levels or the upper and lower bounds of the problem set. In this respect this approach gives a customized methodology to a particular scheduling problem set.

III. AREAS OF FUTURE RESEARCH

This hybridized approach of NaACO has been applied and tested on a simple 3 work stations, 12 jobs environment. It is recommended to test this approach on more complex problems to check the robustness of the model and to recommend changes where necessary. Moreover, in the composition of this model the sensitivity factors of τ and η are taken as constant(i.e. $\alpha,\beta = 1$). In this respect a sensitivity analysis or post optimal analysis is the area of immediate future research for NaACO technique.

IV. CONCLUSION

The essence of this paper is to develop a neural based self-learning and evaluating mechanism within the basic framework of ACO. In this regards a simple scheduling problem is picked and the approach is tested to evaluate the results. The intelligent pheromone up-gradation mechanism is recommended through the inclusion of ANN in the pheromone update mechanism.

REFERENCES

[1] M. Dorigoand S.Thomas, "Ant Colony Optimization." London: The MIT press, 2004.

[2] M. Umer, Dr. R. Ahmad, Dr. I. Chaudhry, "Unsupervised Artificial Neural Networks (ANNs) For Intelligent Pheromone up Gradation. FurtherEvolution of Neural Augmented Ant Colony Optimization (NaACO)," *Life Science* 10 ed. vol. 3, 2013, pp. 318-327.

ISBN: 978-1-61804-227-9

[3] M. R. Garey and D. S. Johson, *Computer and Intractability: A Guide to the theory of NP-Completeness*, 1979.

[4]M. Mathirajan,V. Chandru, A. I. Sivakumar, "Heuristic algorithms for scheduling heat-treatment furnaces of steel casting industries." *Sādhanā*32ed. vol. 5, 2007, pp. 479-500

[5] P. P. Grasse, "La reconstruction du nid et les coordinationsinterindividuelleslles chez. bellicositermesnatalensis et cubitermessp 6," 1959, pp. 41-81.

[6] M. Dorigo, "Optimization, Learning and Natural Algorithms." PhD thesis, Politecnico di Milano, Dipartimento di Eletronica, Milan.1992.

[7] M. L den Besten, T.Stutzle, and M. Dorigo, "Ant colony optimization for the total weighted tardiness problem,"*in Proc. of PPSN-VI, Sixth International Conference on Parallel Problem Solving from Nature*, Berlin: Springer-Verlag. 2000, pp. 611-620.

[8] C. Rajendra, and H. Ziegler, "Ant-colony algorithms for permutation flowshop scheduling to mini-mize makespan/total flow time of jobs." *European journal of Operational Reseacrh*, 2003.

[9] D. Merkle,M. Midderdorf, and H. Schmeck, "Ant colony optimization for resource-constrained project scheduling."*IEEE Transactions on Evolutionary Computation*6 ed. vol. 4, 2002, pp.333-346.

[10] C. Blum and M. Sampels, "Ant colony optimization for FOP shop scheduling: A case study on different pheromone representations,"*in Proc. of the 2002 Congress on Evolutionary Computation (CEC-02)*, Piscataway, NJ: IEEE Press, 2000a, pp-1558-1563.

[11] D. E. Rumelhart, and J. L. McClelland, *Parallel Distributed Processing*. Cambridge, MA: MIT Press.1986.

[12] J. Hopfield and D. W. Tank, "Neural computation of decisions in optimization problems." *Biological Cybernetics*52 ed., 1985.

[13] M. O. Poliac, E. B. Lee, J. R. Slagle, and M. R. Wick, "A Crew Scheduling Problem."*in Proc.of IEEE International Conference on Neural Networks*, 1987, pp. 779-786.

Applying Fuzzy Delphi Method and Fuzzy Analytic Hierarchy Process for Ranking Marine Casualties

Akbar Etebarian[1] , Alireza Shirvani[2] , Iraj Soltani[3] Ali Moradi[4]
Department of Public Management Science and Research Azad University , Isfahan Branch, Shahraz Alley 1st Abshar St. Isfahan Iran
1.Etebarian@ymail.com 2. Baleandish@yahoo.com 3. I.soltani@msc.ir 4. Ali8ir@Yahoo.com (Responsible Author) ,

Abstract: Marine accidents, particularly those that involve pollution and large fatalities, bring into question the safety of shipping and the quality of ships and their crews. Whether or not such questions are justified, it is marine accidents that provide a poor image of the industry, which attract considerable attention. Incidents that particularly attract attention are those causing loss of life, pollution of the environment and the loss of ship and or cargo. Usually, People have a tendency to focus on the consequences of an accident rather than its root causes, so MCDM[1] could improve to find the root cause elements by providing more precise decision parameters. Due to the complexity of Marine accident investigation, this study aims to provide a systematic approach to determine the degree of most influence parameters (cause and effect) in accident occurrence, in order to improve marine safety in direction of Good Governance; in the study two phase procedures are proposed. The first stage utilizes Fuzzy Delphi Method to obtain the critical factors of the Marine Accident Investigation by interviewing the related connoisseurs. In the second stage, Fuzzy Analytic Hierarchy Process is applied to pair fuzzy numbers as the measurable indices and finally to rank degree of each influence criterion within accident investigation. This study considers 1 Goal, 4 Aspects and 31 Criteria (Parameters), and establishes a ranking model that provides decision makers to assess the prior ordering of reasons and sorts by most effective parameter involved Marine Accident occurrence. The empirical study indicates that the " People, Working and living conditions, Effect " is the high ranking aspect and "Ability, Skills and knowledge of workers" is the most important evaluation criterion considered in overall experts view derived from Fuzzy Delphi Analytical Hierarchy Processing (FDAHP). The demonstration of how the prior order of accident maker parameters of connoisseurs is addressed as well. Therefore, ranking the priority of every influencing criterion (parameter), shall help the decision makers in marine transportation, to emphasize the area to improve and act accordingly to prevent future marine accidents.

Key-Words: Marine Accident, Accident Investigation, Good Governance, Analytic Hierarchy Process (AHP), Fuzzy Delphi Method (FDM), Fuzzy Analytic Hierarchy Process (FAHP),

1 Introduction

The efficient Marine Accidents Investigation is aimed to prevent the occurrence of same accident in the future. Accidents are rarely simple and almost never result from a single cause. Most accidents involve multiple, interrelated causal factors. Accidents can occur whenever significant deficiencies, oversights, errors, omissions, or unanticipated changes are present. Any one of these conditions can be a precursor for an accident; the only uncertainties are when the accident will occur and how severe its consequences will be. To conduct a complete accident investigation, the Parameters (factors) contributing to an accident, must be clearly understood. Management prevents or mitigates accidents by identifying and implementing the appropriate controls and barriers.

Parameters those involve in marine accident are as follow but not limited to them:

(A) Controls (e.g. rules/regulations, procedures, training, etc.) help to prevent errors or failures that could result in an accident;
(B) Barriers (e.g. emergency systems onboard, contingency plans) help to mitigate the consequences of potential errors or failures. Barriers to protect targets against loss can be physical barriers, such as machine guards and railings, or management barriers, such as work procedures, hazard analysis, requirements management, line management oversight, and communications.

In a work environment, several barriers may be used in an effort to prevent accidents. Proper Marine Accident Analyzing could help reduce both the human loss and environmental pollution from ships, thus, creating a huge efficiency either from environmentally- friendly or economic levels.

This study tries to identify the influencing criteria by using expert's view who have wide knowledge and have been involved in Marine Accident Investigation in Iran.

According the SOLAS (74)[2] convention, each contracting government is responsible to carry proper Maine Accident Investigation process and keep recording for further use. The aim is to get advantage of worldwide Marine Accident Investigation findings to prevent future accidents and lessons can be learnt from each individual inspection.

So ranking the causal elements or reasons of a marine accidence will not only contribute in Good Governance process, but also to environmental protection.

This article introduces a model which takes advantage of expert idea to rank each parameter which are playing roles in the marine accident occurrence process. At present, the most important model for analyzing marine accident causations, which is recommended by IMO[3], is SHEL[4] model which analyzes the software, hardware, environment and life ware, surrounding a marine accident. But the model is not ranking the degree of effective of each parameter. FDAHP approach provides a scientific decision making in marine accident investigation by providing reliable decision factors for decision makers.

[1] . Multiple Criteria Decision Making

2 .Safety Of Life At Sea Convention 1974 as amended

3 . International Maritime Organization (IMO)

4 .Software , Hardware, Environment, Life ware (SHEL)

ISBN: 978-1-61804-227-9

2. Literature Reviw

2.1 Good Governance

Collely, John L, Doyle, Jacqueline L,(2003) and Osborne, (2009), mentioned the characteristics of Good Governance, which is now considered as a management paradigm , as follows:

A. ***Accountability*****:** is the key for good governance. Decision makers are accountable to those who are affected by the decision and implementation. Accountability cannot be enforced without transparency and the rule of law.

B. Responsiveness: is the concerns of those who are affected, those who implement, and those who integrate in the formulation of the decision.. To be responsive is also to provide feed back and address grievances spontaneously,

C. Transparency: means that the information is freely and directly accessible to those affected by the decision. It also means that the decisions are taken and enforced strictly within the established rules and regulations,

D. ***Citizenship Satisfaction*****:** this is most vital character and is the degree of satisfaction of people involved.

Therefore the MCDM is scientific approach to implement Good Governance as this study tries to high light it.

2.2 Marine Accident Investigation

2.2.1 Responsibilities of States

Each State shall cause an inquiry to be held by or before a suitably qualified person into every casualty or incident of navigation on the high seas involving a ship flying its flag and causing loss of life or serious injury to nationals of another State or serious damage to ships or installations or another State or to the marine environment. The flag State and the other State shall co-operate in the conduct of any inquiry held by other State into any such marine casualty or incident of navigation[5].

The objective of any marine casualty investigation is to prevent similar casualties in the future.

Investigations identify the circumstances of the casualty under investigation and establish the causes and Contributing factors, by gathering and analyzing information and drawing conclusions. Ideally, it is not the purpose of such investigations to determine liability, or apportion blame. However, the investigating authority should not refrain from fully reporting the causes because fault or liability may be inferred from the findings.

Over the years, and as a result of some major accidents, some of the existing international instruments have changed and some others specifically created to deal with various aspects of marine casualties. The most important ones are mentioned here and one in particular, (IMO Casualty Investigation Code), that is central to this course, will be covered in the next lesson. These are:

a) UN Convention on the Law of the Sea (UNCLOS)
b) IMO Conventions
c) IMO Assembly Resolutions
d) IMO MSC Circulars and Codes
e) International Labor Organization (ILO) Conventions

Marine Casualty Investigation Authorities' main objective of any marine casualty investigation is prevention of further similar cases by discovering the reasons behind the casualty and then promulgating actions, information and recommendations where appropriate, with a view to preventing similar casualties. Other benefits and reasons for investigations include:

1) Improved design
2) Improved operational and safety procedures
3) Improved work environment
4) Improved safety awareness

It is important that any recommendation arising from an investigation is based on sound analysis and is capable of practical implementation.

It follows from this that any casualty, from the simple to the major, can be the subject of a marine casualty investigation. A simple personnel incident, with the potential for learning something which could prevent recurrences, might be worth investigating thoroughly while a major collision resulting from a straightforward wrong application of the COLREGS[6] might not show anything new. Another different collision might indicate a need to look at fatigue, management procedures, training, certification and bridge design. The depth that each casualty which is reported needs to be investigated should be assessed on its merits.

2.2.2 What is a Marine Causality or Accident?

A marine Causality or Accident can be considered:

Marine casualty means an event that has resulted in any of the following:

a) the death of, or serious injury to, a person that is caused by, or in connection with, the
b) operations of a ship; or
c) the loss of a person from a ship that is caused by, or in connection with, the operations of a
    a. Ship; or
d) the loss, presumed loss or abandonment of a ship; or
e) material damage to a ship; or
f) the stranding or disabling of a ship, or the involvement of a ship in a collision; or
g) material damage being caused by, or in connection with, the operation of a ship; or
h) damage to the environment brought about by the damage of a ship or ships being caused by,
i) Or in connection with, the operations of a ship or ships. (IMO Resolution A.849 (20) adopted on 27 November 1997)

It was reported that 61 seafarers lost their lives on commercial vessels operating in and around EU waters in 2010 (compared with 52 in 2009 and 82 in both 2008 and 2007). The majority of these were in accidents involving fishing vessels (33%), while accidents on general cargo ships accounted for 28% of lives lost in 2010 (European Maritime Safety Agency Maritime Accident Review 2010,)

According to the Iranian maritime authorities only more than 100 vessels had accident during 2012 losing life, environmental impact and ship and cargo damages.(www.pmo.ir)

5 . Article 94, Duties of the flag State, provides, in paragraph 7 , United Nation Convention on Law Of Sea)

6 . Collision Regulation at Sea

ISBN: 978-1-61804-227-9

2.3 Historical Overview of accident investigation

The sinking of the passenger liner SS[7] "Titanic" in 1912 made shipping safety a matter of public concern and issue, which later led to the development of the first SOLAS Convention in 1929 and formation of an international organization responsible for the safety of international shipping, now known as the International Maritime Organization (IMO). After then it was great importance to evaluate the reasons of each marine incident. IMO has provided codes and guidelines for effective marine accident investigation and data bank to collect information about accidents worldwide. So the marine accident investigation came to attention of all marine community especially governmental authorities.

3. Validity and Reliability

Warwick and Linninger (1975) point out that there are two basic goals in questionnaire design.

1. To obtain information relevant to the purposes of the survey.

2. To collect this information with maximal reliability and validity.

How can a researcher be sure that the data gathering instrument being used will measure what it is supposed to measure and will do this in a consistent manner? This is a question that can only be answered by examining the definitions for and methods of establishing the validity and reliability of a research instrument. These two very important aspects of research design will be discussed in this module.

3.1 Validity

Validity can be defined as the degree to which a test measures what it is supposed to measure. There are three basic approaches to the validity of tests and measures as shown by Mason and Bramble (1989). The validity for questionnaire is obtained by **KMO and Bartlett's Test** by SPSS19 software.

3.2 Reliability

The reliability of a research instrument concerns the extent to which the instrument yields the same results on repeated trials. Although unreliability is always present to a certain extent, there will generally be a good deal of consistency in the results of a quality instrument gathered at different times. The tendency toward consistency found in repeated measurements is referred to as reliability (Carmines & Zeller, 1979). In scientific research, accuracy in measurement is of great importance. Scientific research normally measures physical attributes which can easily be assigned a precise value. Many times numerical assessments of the mental attributes of human beings are accepted as readily as numerical assessments of their physical attributes. Although we may understand that the values assigned to mental attributes can never be completely precise, the imprecision is often looked upon as being too small to be of any practical concern. However, the magnitude of the imprecision is much greater in the measurement of mental attributes than in that of physical attributes. This fact makes it very important that the researcher in the social sciences and humanities determine the reliability of the data gathering instrument to be used (Willmott & Nuttall, 1975). Reliability of questionnaire is obtained by Cronbach's Alpha Test by SPSS19 software.

7 . Steam Ship

4. Methodology

The study contains two stages: the first stage is to establish the key parameters for evaluation of the marine accident analyzing, and use FDM by consulting experts from government sectors, academia and shipping industries to select a criterion, in order to find out the important factors to be conceded. We selected four organizations which are involved in marine activities, namely; PMO[8], IRISL[9], NIOTC[10]and two Universities[11]. The second stage is based on FAHP, and consults high level experts of various sections to find out the importance of various criteria, in order to obtain the measuring index for selecting the effective degree of each parameter on a marine accident occurrence. The survey methodology was used to gather the data and to build the marine accident causal criteria. Before designing the survey, we gathered the evaluation criteria from literature studies and some expert idea. Beside, according the literatures, we combined the criteria of accident causal elements and prior researches in related or other arenas, and generalized 43 factors of which 31 selected as important constructs under four important aspects.

4.1. Fuzzy Delphi Method

Fuzzy Delphi Method was proposed by Ishikawa et al. (1993), and it was derived from the traditional Delphi technique and fuzzy set theory. Noorderhaben (1995) indicated that applying the Fuzzy Delphi Method to group decision can solve the fuzziness of common understanding of expert opinions. As for the selection of fuzzy membership functions, previous researches were usually based on triangular fuzzy number, trapezoidal fuzzy number and Gaussian fuzzy number. This study applied the triangular membership functions and the fuzzy theory to solving the group decision. This study used FDM for the screening of alternate factors of the first stage. The fuzziness of common understanding of experts could be solved by using the fuzzy theory, and evaluated on a more flexible scale. The efficiency and quality of questionnaires could be improved. Thus, more objective evaluation factors could be screened through the statistical results.

The FDM steps are as follows:

1 . *Collect opinions of decision group*: Find the evaluation score of each alternate factor's significance given by each expert by using linguistic variables in questionnaires.

2 . *Set up triangular fuzzy numbers:* Calculate the evaluation value of triangular fuzzy number of each alternate factor given by experts, find out the significance triangular fuzzy number of the alternate factor. This study used the geometric mean model of mean general model proposed by Klir and Yuan (1995) for FDM to find out the common understanding of group decision. The computing formula is illustrated as follows:

Assuming the evaluation value of the significance of No. j element given by No. i expert of n experts is w_{ij} (a_{ij} , b_{ij} , c_{ij}), i =1, 2, ..., n; j =1, 2, ..., m. Then the fuzzy weighting w_j of No. j element is w_j (a_j , b_j , c_j), j =1,2, ..., m

8 . Ports and Maritime Organization (Maritime Authority in Iran)
9 . Islamic Republic of Iran Shipping Line
10 . National Iranian Oil Tanker Company
11 . Chabahar Nautical University and Khoramshar marine science and Technology University.

ISBN: 978-1-61804-227-9

$a_j = Min_i(a_{ij})$, $b_j = \frac{1}{n}\sum_{i=1}^{n} bij$, $c_j = Max_i(c_{ij})$

3. Defuzzification: Use simple centre of gravity method to defuzzify the fuzzy weight w_j of each alternate element to definite value S_j , the followings are obtained:

$S_j = \frac{aj+bj+cj}{3}$ j=1, 2, ..., m

4. Screen evaluation indexes: Finally proper factors can be screened out from numerous factors by setting the threshold a. The principle of screening is as follows:

If $S_j \geq \alpha$, then No. j factor is the evaluation index.

If $S_j < \alpha$, then delete No. j factor. Schematic diagram of Fuzzy Delphi Method threshold is shown in Fig. 1.

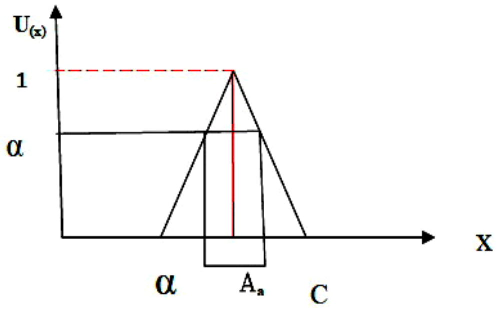


Fig. 1. Schematic diagram of Fuzzy Delphi Method threshold

4.2. Fuzzy Analytic Hierarchy Process

Laarhoven and Pedrycz (1983) proposed the Fuzzy Analytic Hierarchy Process in 1983, which was an application of the combination of Analytic Hierarchy Process (AHP) and Fuzzy Theory. The linguistic scale of traditional AHP method could express the fuzzy uncertainty when a decision maker is making a decision. Therefore, FAHP converts the opinions of experts from previous definite values to fuzzy numbers and membership functions, presents triangular fuzzy numbers in paired comparison of matrices to develop FAHP, thus the opinions of experts, approach human thinking model, therefore as to achieve more reasonable evaluation criteria.

As for the experts' opinions, this study adopted the Similarity Aggregation Method (SAM) proposed by Hsu and Chen (1996)[12] to integrate experts' weight values for various evaluation criteria, the fuzzy weight fraction of criterion of each hierarchy is obtained through the calculating mode of FAHP, and then the sequence of significance of each criterion is determined based on the hierarchy series connection and defuzzification mode.

Laarhoven and Pedrycz (1983) proposed the FAHP, which is to show that many concepts in the real world have fuzziness. Therefore, the opinions of decision makers are converted from previous definite values to fuzzy numbers and membership numbers in FAHP, so as to present in FAHP matrix.

The steps of this study based on FAHP method are as follows:

1.*Determine problems:* Determine the current decision problems to be solved, so as to ensure future analyses correct; this study discussed the ‘‘evaluation criteria for Marine Accident Investigation”.

2 . *Set up hierarchy architecture:* Determine the evaluation criteria having indexes to be the criteria layer of FAHP, for the selection of evaluation criteria, relevant criteria and feasible schemes can be found out through reading literatures and collective discussions. This study screened the important factors conforming to target problems through FDM investigating experts' opinions, to set up the hierarchy architecture (as shown in fig. 3).

3 . *Set up fuzzy paired comparison matrices:* Compare the relative importance between factors given by decision makers in pairs, set up paired comparison matrices, after the definite values are converted to fuzzy numbers according to the definition in Table 1 and Fig. 2, integrate the fuzzy evaluation values of experts based on the Similarity Aggregation Method SAM concept proposed by Hsu and Chen (1996).

4 .*Calculate fuzzy weight value*: Obtain the characteristic vector value of fuzzy matrix, namely the weight value of element. This study calculated these three positive and negative value matrices respectively by using the ‘‘Column Vector Geometric Mean Method” proposed by Buckley (1985).

5. Hierarchy series connection: Connect all hierarchies in series, to obtain all factors' weights.

To collect the fuzzy numbers which have derived from directly from expert idea, In this study we have

used triangular method therefore a fuzzy number s defined according relations numbers (1) to (4):

(1) $\alpha_{ij} = (\alpha_{ij}, \mathbf{d}_{ij}, \mathbf{g}_{ij})$

(2) $\alpha_{ij} = Min(b_{ijk})$, k=1,….n

(3) $\mathbf{d}_{ij} = (\prod_{k=1}^{n} b_{ijk})^{1/n}$, k=1,….,n

(4) $\mathbf{g}_{ij} = Max(b_{ijk})$, k=1,….,n

Fig.2 shows a typical fuzzy number which we have used in this study:

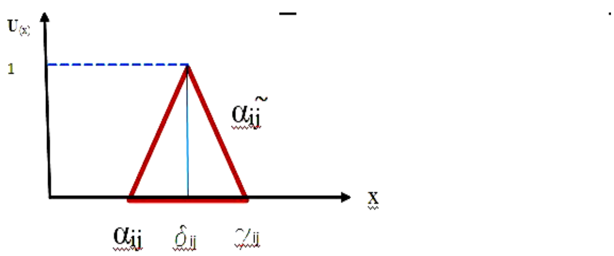


Fig. 2. Schematic diagram of Fuzzy Delphi Method threshold (*Liu and Chen, 2007*)

In which b_{ijk} is the relative preference parameter “i” to parameter “j” from expert “k” view, α_{ij} and g_{ij} are the lower and upper limits of expert view, respectively and d_{ij} is the geometric mean of experts views. Therefore parameters are so defined that: $\alpha_{ij} \leq d_{ij} \leq g_{ij}$

Then according to calculated fuzzy numbers as mentioned above, paired matrices between various parameters the inverted matrices are set up for fuzzy numbers according relation (5):

[12] . A similarity aggregation method (SAM) aggregates experts' opinions in a linguistic framework using a consensus weight factor for each expert that is based on the similarity of his or her opinion relative to the other experts to ensure that the experts' final decision is a result of common agreement Read More: http://ascelibrary.org/action/showAbstract?page=432

ISBN: 978-1-61804-227-9

(5) $A_{ij}=[\alpha_{ij}], \alpha_{ij\times} \alpha_{ji} \approx 1, \forall i,j=1,2,3.....$

To calculate the fuzzy relative weight we have used the following relations numbers (6), (7) and (8):

(6) $Z=[\alpha_{ij} \otimes .... \otimes \alpha_{ij}]$
(7) $Z_i=[\alpha_{ij} \otimes .... \otimes \alpha_{ij}]^{-1}$
(8) $W_{i=} Zi \otimes (Z_i \oplus .... \oplus Z_n)$

6. Defuzzification: Convert fuzzy numbers to easy-comprehended definite values, this study adopts the geometric mean method to solve fuzzy numbers proposed by *Liu and Chen, (2007), according to relation number :*

2. (9) $W_i=$

7. Sequencing: Sequence defuzzified criteria.

Table 1

The definition of every fuzzy number

Fuzzy number Definition	
1= (1,1,1)	Equally important
3=(2,3,4)	Moderately more important
5=(4,5,6)	Strongly more important
7=(6,7,8)	Very strongly more important
9=(8,9,9)	Extremely important

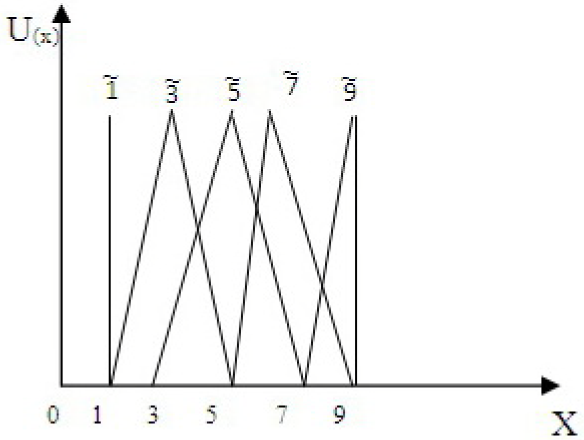


Fig.3. Scale of fuzzy numbers

5. Evaluating model application and results

A) *Reviewing relevant literature of Marine accident Investigation and proposing important criteria*:

More than 43 criteria (Parameter) for Marine Accident Investigation based on reviewing relevant literature (*Liu and Chen,2007*, Begum, Siwar, Pereira, & Jaafar, 2006; Emery, Davies, Grffiths, & Williams, 2007; Finn -veden, 1999; International Maritime Organization (IMO) Resolution A.849(20) dopted on 27 November 1997 ,IMO Resolution A.884(21) adopted on 25 November 1999, Lin, Lin, & Jong,2007) and the current Marine Accident Investigation approach are proposed. A brief definition of evaluating criteria of Marine Accident Investigation is presented in Table 2.

B) *Screen important criteria (Parameters) by Fuzzy Delphi Method*. This stage includes three sections. Firstly, it lists Four main aspects and 43 items as the key evaluation items of Marine Accident Investigation, and a FDM interview framework is set up.The second section is the interview with twenty experts from national shipping company, the academic community, and competent government authority officers in Iran. Delphi Method mostly aims at easy common understanding of group opinions through twice provision of questionnaires. FDM formed by adding the fuzzy theory in, not only maintains the advantage of Delphi Method, but also reduces the provision times of questionnaires when using traditional Delphi Method as well as the cost.

For the third section, the opinions of experts in FDM questionnaires are converted to triangular fuzzy numbers, and defuzzified values can be figured out after calculation. This stage adopts elements with threshold above 6, and the key evaluation items with threshold below 6 are deleted. The important evaluation items after screening are listed in Table 3.

ISBN: 978-1-61804-227-9

Table 2
Operational type for defining for 43 criteria

Aspects	Criteria (Parameters)	Short Operational Definition
Organization on board & Shore-side management (A)	Division of tasks and Responsibilities	Written job description, task analyzing, responsibility allocation, ...
	Composition of the crew	Mixture of nationality of crew and their competence and training.
	Working hours (planned)	Schedule duty, day or night time worker...
	Workload / Complexity of tasks	Amount of Paper work, Bureaucratic Activities etc.
	Rest hours (planned)	Including sleep duration and time for recreation.
	Procedures and standing orders	The way to implement the written current orders on board.
	Communication	Internal and external communication procedures.
	On-board management and supervision	Written mechanisms which make sure the works are going right direction.
	Organization of on-board training	Organizing practical on-board training to updated the workers.
	Organization of on-board drills	Written Procedures to carry out drills on-board.
	Voyage, cargo and maintenance ,planning	The procedure which a voyage or cargo or maintenance and etc, are planned.
	Policy on Recruitment	Written procedure on how the company select workers.
	Safety Policy and Philosophy	Written safety policy and training procedures including emergency drills
	Management Commitment to Safety	Written procedure from high level management indicating safety commitment.
	Amount of Logistic Support from Shore	Written policy of organizational logistic support
	Policy for workers motivation	Written policy of Management procedure for motivating the workers.
	Port Scheduling	Planning to leave or arrive on a port, stay at port...
	Contractual arrangements	Contractual, industrial arrangements and agreements for all crew members.
	Assignment of Duties	Assigning the duties to the involved workers.
	Ship-Shore Communication	Interaction with port, headquarter, emergency stations etc.
Ship factors (B)	Design of Ship and Equipment	Quality of Ship and Equipments Design....
	State of Maintenance	The Condition Which The Equipment Is Maintained...
	Equipment	The Availability, Reliability, Durability Performance Of Equipments….
	Cargo characteristics	Including Securing, Handling And Care Of Cargo ….
	Certificates	Certificates For Ship, Equipment, Machinery, …
	Ship type	Including Cargo Ship, Crude Oil Carrier, Ro-Ro Ship, Passenger Ship, …
Environment (C)	Weather and Sea Conditions	Internal And External Climate, Temperature. Visibility, Vibration, Noise ...
	Port and Transit Conditions	Including Vessel Traffic Service , Pilots, Port Facilities, Etc.
	Traffic Density	Number Of Coming And Going Vessels In The Area.
	Heavy Weather Conditions	Wind, Rain, Snow, Tifton, Cyclone...
	Representing Agencies	The Ship Owners And Seafarers Representatives And Agencies.
	Regulations, Surveys and Inspections	International, National, Port, Classification Societies, Etc.
	Shore side Interaction	Interaction With Stevedores, Port Officials, Security Measure In Port Area Etc.
People & Working and living conditions (D)	Ability, Skills, knowledge	The Outcome Of Training, Experience, Education, Professional, Certification…
	Personality	The Mental Condition, Emotional State,
	Physical condition	Including Sickness, Medical Fitness, Drugs And Alcohol, Fatigue…
	Sleep and its quality	Scheduled Sleep And The Area Which The Sleep Takes Place.
	Person Abilities	Assigned Duties Respect To Person Abilities
	Actual behavior at time of Accident	The Location, Task Performing, Attention, At Time Of Accident Occurrence
	Level of Automation	Taking Advantage Of Automatic Instruments To Perform Tasks And Duties.
	Ergonomic Design	Working, Living And Recreation Areas And Equipment Suitable For Human.
	Adequacy of Living Conditions	Opportunities For Recreation, Rest, Sleep…
	Adequacy of Food	The Quality And Quantity Of Food For Workers To Carry Out Their Duty.

ISBN: 978-1-61804-227-9

Table 3
New evolution Criteria after Fuzzy Delphi Method

Aspects	Criteria(Parameters)- Code Number		Score Min	Mean	Max	Defuzzification
Organization on board & Shore-side management (A)	Division of tasks and Responsibilities	A_1	5	7.4	9	7.1
	Working hours (planned)	A_2	5	6.5	9	6.8
	Workload / Complexity of tasks	A_3	3	6.4	9	6.1
	Rest hours (planned)	A_4	3	7.2	9	6.4
	Procedures and standing orders	A_5	3	6.2	9	6.1
	Communication	A_6	3	6.3	9	6.1
	On-board management and supervision	A_7	3	6.3	9	6.1
	Organization of on-board training	A_8	3	6.4	9	6.1
	Organization of on-board drills	A_9	3	7	9	6.3
	Policy on Recruitment	A_{10}	3	6.4	9	6.1
	Management Commitment to Safety	A_{11}	3	7.3	9	6.4
	Amount of Logistic Support from Shore	A_{12}	3	6.7	9	6.2
	Policy for workers motivation	A_{13}	3	6.4	9	6.1
	Assignment of Duties	A_{14}	3	6.7	9	6.2
Ship factors (B)	Design of Ship and Equipment	B_1	5	7	9	7.0
	State of Maintenance	B_2	5	7.4	9	7.1
	Equipment	B_3	5	7.2	9	7.1
	Cargo characteristics	B_4	3	6.5	9	6.2
Environment (C)	Weather and Sea Conditions	C_1	5	6.7	9	6.9
	Port and Transit Conditions	C_2	5	6.8	9	6.9
	Traffic Density	C_3	5	6.2	9	6.7
	Heavy Weather Conditions	C_4	5	6.6	9	6.9
	Regulations, Surveys and Inspections	C_5	5	6.5	9	6.8
People & Working and living conditions (D)	Ability, Skills, knowledge	D_1	7	8.4	9	8.1
	Personality	D_2	5	7	9	7.0
	Physical condition	D_3	5	7.1	9	7.0
	Sleep and its quality	D_4	5	7.6	9	7.2
	Person Abilities	D_5	5	7.5	9	7.2
	Actual behavior at time of Accident	D_6	5	7.5	9	7.2
	Ergonomic Design	D_7	5	6	9	6.7
	Adequacy of Living Conditions	D_8	3	6.5	9	6.2

C) *Establish a hierarchical framework:*

Based on the FDM, a general consensus among experts can be reached to establish a hierarchical structure. The Marine Accident Investigation can be evaluated based on four evaluation aspects and 31 evaluation criteria or Parameters (Fig. 2).

D) *Interview experts and integrate their opinions:* Subject to who fill in AHP questionnaires possess sufficient professional knowledge and at least 20 years of experience in marine activities either in shipping or authorising (government) field , so the interviewees are experts and from different concerned activities. The evaluation of each factor must go through consistency verification to ensure preferable credibility of results. In order to increase the objectivity of results, there are twenty experts to be interviewed. In the past, the integration of opinions from questionnaires adopted geometric mean method, but the unreasonable integration of group opinions therein would result in incorrect results. Therefore, this study adopts Similarity Aggregation Method (SAM) which was proposed by Hsu and Chen (1996), which can integrate group opinions more reasonably, so as to increase the credibility of questionnaires.

E) *Calculate the weights of evaluation criteria and weight result of evaluation criteria:* The weight values of various elements can be obtained through the opinions of experts resulted from SAM and the FAHP systematic steps. After sequencing, the evaluation criteria have higher significance, so decision makers can make correct judgments more quickly.

ISBN: 978-1-61804-227-9

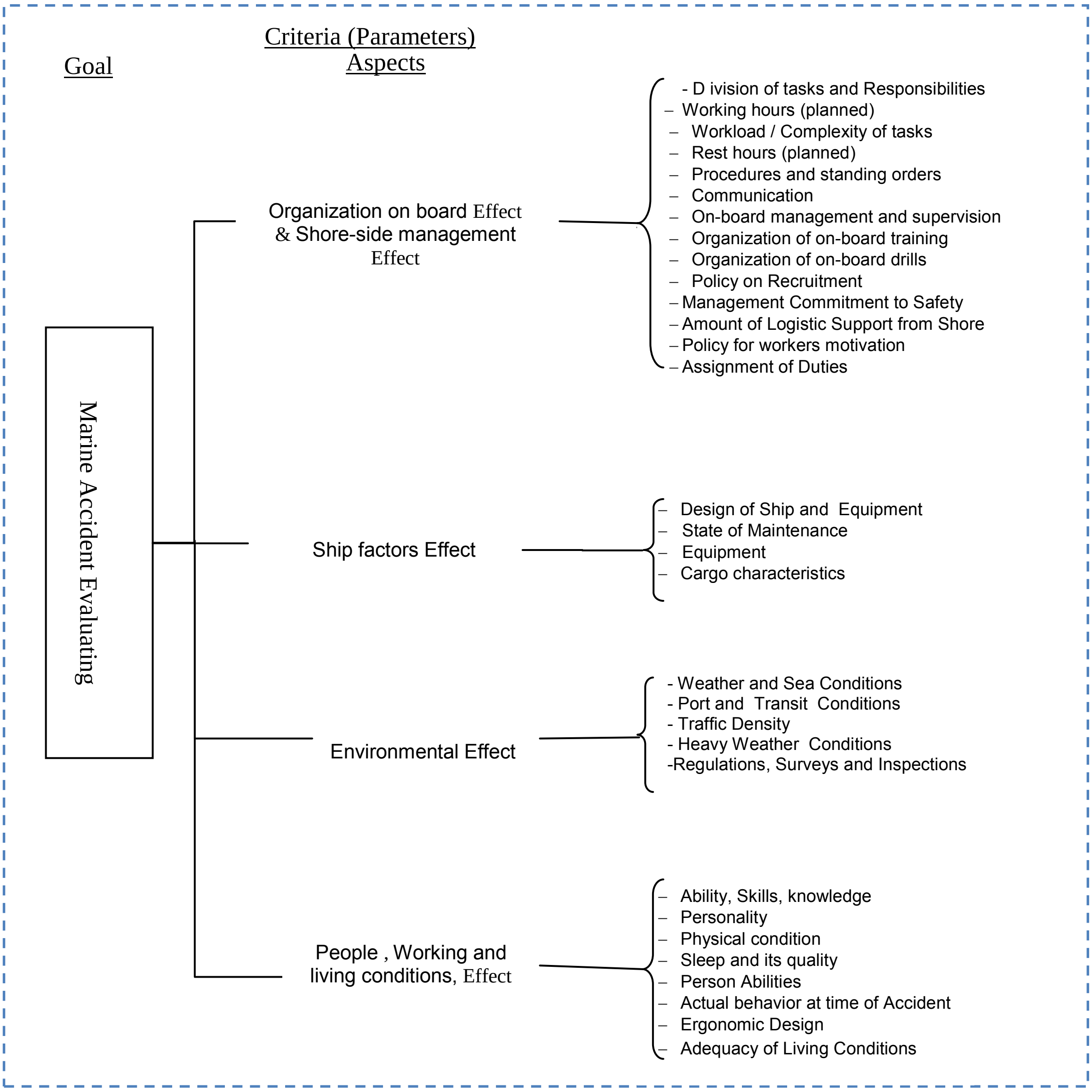


Fig. 4. The Hierarchy model of Marine Accident Investigation

Table 4 is the evaluation criteria weight by FAHP, the evaluation criteria weight is obtained based on FAHP questionnaire results of experts, finally the questionnaire results of all experts are integrated to become the overall weight.

Table 4
Evaluation Criteria Weight of connoisseurs

Aspects	Weights of Aspects	Criteria	Priority	Weights of Criteria: Fuzzy Weights W_i			Defuzzification	Sort largest to smallest	Ranking	
(A)	0.253	A_1	6	0.013	0.0353	0.1164	0.0376	0.0444	D_1	1
		A_2	16	0.012	0.0310	0.0997	0.0337	0.0389	D_4	2
		A_3	25	0.009	0.0297	0.1081	0.0306	0.0388	D_5	3
		A_4	14	0.011	0.0335	0.1167	0.0346	0.0380	B_2	4
		A_5	24	0.010	0.0291	0.0999	0.0308	0.0376	D_6	5
		A_6	29	0.009	0.0295	0.1008	0.0298	0.0376	A_1	6
		A_7	31	0.008	0.0294	0.0944	0.0287	0.0371	B_3	7
		A_8	26	0.009	0.0302	0.1000	0.0304	0.0361	B_1	8
		A_9	19	0.010	0.0330	0.1061	0.0330	0.0359	D_2	9
		A_{10}	30	0.009	0.0300	0.0942	0.0295	0.0359	D_3	10
		A_{11}	17	0.010	0.0342	0.1101	0.0332	0.0357	C_1	11
		A_{12}	27	0.009	0.0308	0.1034	0.0300	0.0352	C_4	12
		A_{13}	28	0.009	0.0297	0.1004	0.0298	0.0348	C_2	13
		A_{14}	22	0.010	0.0315	0.1010	0.0314	0.0346	A_4	14
(B)	0.230	B_1	8	0.013	0.0332	0.1076	0.0361	0.0343	C_5	15
		B_2	4	0.014	0.0354	0.1111	0.0380	0.0337	A_2	16
		B_3	7	0.013	0.0345	0.1097	0.0371	0.0332	A_{11}	17
		B_4	23	0.010	0.0305	0.1008	0.0310	0.0331	C_3	18
(C)	0.216	C_1	11	0.013	0.0319	0.1115	0.0357	0.0330	A_9	19
		C_2	13	0.012	0.0323	0.1056	0.0348	0.0320	D_7	20
		C_3	18	0.012	0.0294	0.1019	0.0331	0.0316	D_8	21
		C_4	12	0.013	0.0315	0.1100	0.0352	0.0314	A_{14}	22
		C_5	15	0.012	0.0307	0.1101	0.0343	0.0310	B_4	23
(D)	0.301	D_1	1	0.017	0.0405	0.1249	0.0444	0.0308	A_5	24
		D_2	9	0.013	0.0335	0.1030	0.0359	0.0306	A_3	25
		D_3	10	0.013	0.0340	0.1044	0.0359	0.0304	A_8	26
		D_4	2	0.014	0.0363	0.1158	0.0389	0.0300	A_{12}	27
		D_5	3	0.014	0.0355	0.1173	0.0388	0.0298	A_{13}	28
		D_6	5	0.013	0.0359	0.1100	0.0376	0.0298	A_6	29
		D_7	20	0.012	0.0286	0.0966	0.0320	0.0295	A_{10}	30
		D_8	21	0.010	0.0307	0.1005	0.0316	0.0287	A_7	31

For different aspects experts have selected aspect (D) “People , Working and living conditions, Effect” (0.301) , aspect (A) “The Organization on board & Shore-side management Effect” (0.253), aspect (B) . “Ship factors Effect”, aspect (0.230) and aspect “ Environmental Effect” (0.216) respectively.

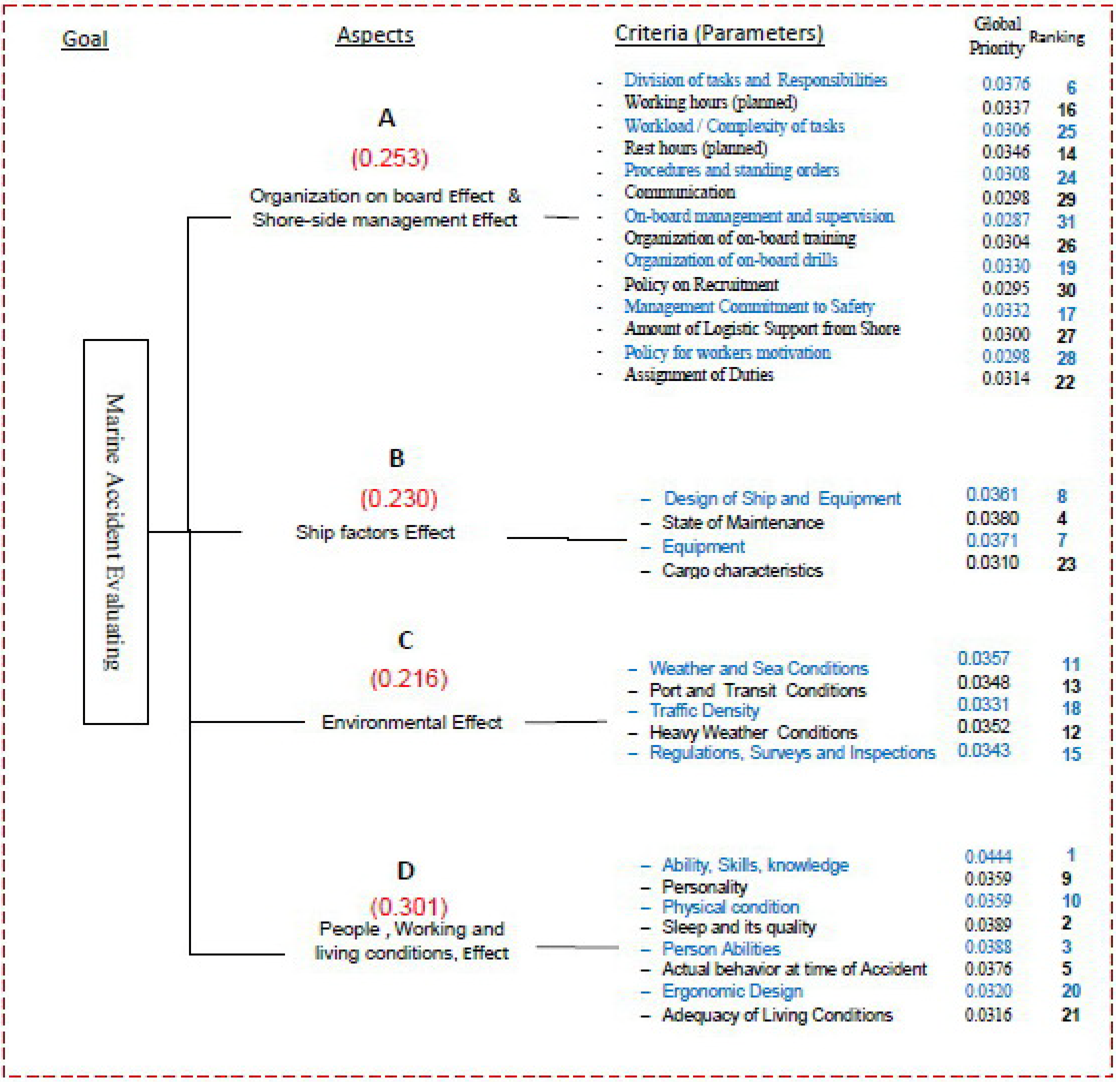


Fig.5. The weights of Marine Accident Investigation Hierarchy Model

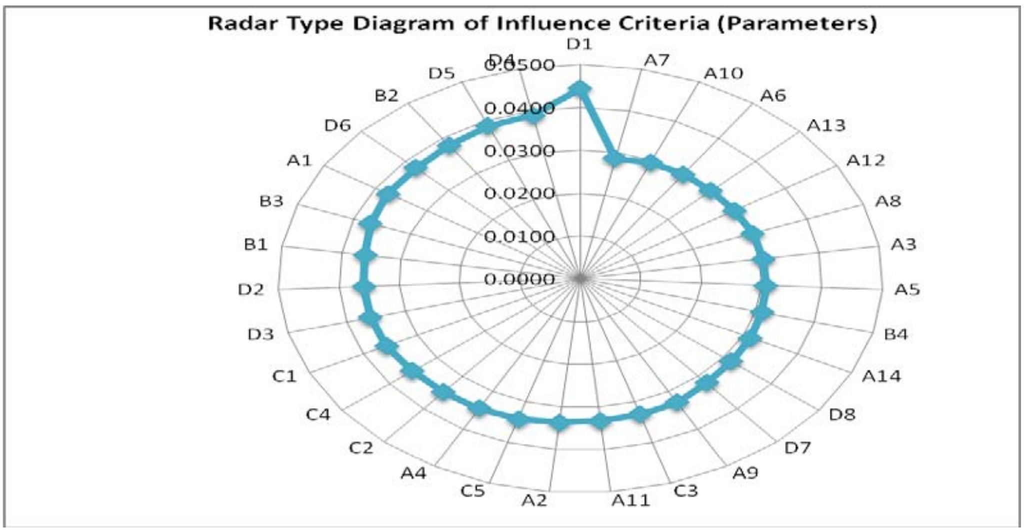


Fig. 6. Most to least, influence parameters (criteria, D_1 to A_7)

6. Conclusions

This study investigates the key factors in marine accident investigation by combining FDM, and FAHP, and establishes objective and standardized references. A total of 43 factors influencing marine accident investigation are analyzed through FDM experts' opinions investigation, Experts of governmental

ISBN: 978-1-61804-227-9

sectors, academia and shipping industry were interviewed, and 31 evaluation criteria were obtained as the key factors (parameters) by interviewed experts. SAM and FAHP were used to integrate experts' opinions to obtain the significance evaluation of various evaluation criteria given by experts in group decision. The results from experts were compared and analyzed to show the similarities and differences of various experts in marine accident investigation. Finally, the results of all experts were used as the evaluation index of marine accident investigation. The following conclusions were reached by analyzing the evaluation criteria stressed by experts when evaluating the marine accident investigation based on the demonstration of this study. The proposed method enables decision analysts to better understand the complete evaluation process. This approach provides a more accurate, effective, and systematic decision support tool.

The importance of the criteria is evaluated by experts, and the uncertainty of human decision making is taken into account through the fuzzy concept in fuzzy environment. From fuzzy AHP we find out that, thirty one criteria out of forty three for marine accident investigation are most important criteria and four aspects, (A) Organization on board Effect & Shore-side management Effect (B) Ship factors Effect (C) Environmental Effect and (D) People, Working and living conditions, Effect, are the most important as shown in Fig. 2.

In this study we highlighted the most important parameters which are assumed to cause a marine accident occurrence in marine accident investigation processing.

1. Emphasis on four main aspects:

The experts have different stress on four aspects; the aspect (D) has a higher weight (0.301), this is probably because the experts include those of People factors namely: (a) ability, skills, knowledge (b) personality (c) physical condition (d) activities prior to accident occurrence (e) assigned duties at time of accident occurrence (f) actual behavior at time of accident occurrence (h) attitude and so on, the experts thought. These are both outcome of training and experience, mental condition, emotional state, medical fitness, use of drugs and alcohol, fatigue etc.

2. Emphasis on over all criteria (five high ranking criteria):

Experts pay highly attention to evaluation criteria in (D)" People, Working and living conditions, Effect" aspect. Therefore the highest ranking criterion is "Ability, Skills, knowledge" (0.0444), the second "Sleep and its quality (0.0389) ", third "Person Abilities" (0.0388) criteria lied in aspect (D) and fifth "Actual behavior at time of Accident" (0.0376) but the forth highest ranking criterion lay in aspect (B) "State of Maintenance" (0.0380) Although the weight of aspect (B) "Ship factors" Effect " (0.230),is third among the four evaluation criteria in that experts make their opinions.

3. Mostly concerned issue:

As we demonstrated above the main parameter in marine accident causation is the criteria "Ability, Skills, knowledge" with weight of (0.0444) in overall expert opinions. Because this is the most important element which directly effect on overall activates. Actually ability, skill and knowledge are the outcomes of training and experience that workers collect over the time. Their definitions are:

ability: power or capacity to do or act physically, mentally, legally, morally, financially, competence in an activity or occupation because of one's skill, training, or other qualification,

Skill: the ability, coming from one's knowledge, practice, aptitude, etc., to do something well: Carpentry was one of his many skills, competent excellence in performance; expertness; dexterity and,

Knowledge: acquaintance with facts, truths, or principles, as from study or investigation; general erudition.

As we understand these are potential factors in every human activity in any field especially in marine transportation. So we suggest the decision makers to consider while planning, organizing, directing, training ..., and people for marine occupations.

7. References

[1]American Bureau of Shipping. (2004). ABS Review and Analysis of Accident Databases: 1991 – 2002 technical Report Number 2003-01).

[2]Article 94,Duties of the flag State, , United Nation Convention on Law Of Sea (UNCLOS 1982)

[3]Ataei M, Fuzzy Multiple Criteria Decision making (FMCDM) (2010) 191-208

[4]Australian Transportation Safety Bureau. (2003). Retrieved Aug, 2003

[5]Cheng, Y. W., Lin, K. H., Chang, K. H., & Huang, W. R. The application of Fuzzy Delphi Method and Fuzzy AHP in lubricant regenerative technology selection, Expert [6]Systems with Applications 37 (2010) 419–425Canadian Transportation Safety Board.(2003).Retrieved July16, 2003

[7]Clifford C. Baker , Ah Kuan Seah , Maritime Accidents and Human Performance: the Statistical Trail American Bureau of Shipping . Presented at MARTECH 2004, Singapore, September 22-24, 2004

[8]Collely,John L, Doyle, Jacqueline L, “Corporate Governance”(Tata McGraw-Hill Publishing Company Ltd., 2003)

Corporate Behaviour (http://en.wikipedia,org/wiki/corporate_behaviour)

[9]Dalkey, N., & Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. Management Science, 9, 458–467.

[10]Houston, TX: American Bureau of Shipping. (Internal SAHF Report).

[11]Hsu, H. M., & Chen, C. T. (1996). Aggregation of fuzzy opinions under group decision making. Fuzzy Sets and System, 79, 279–285.

Http://dictionary.reference.com/browse/knowledge?s=t

[12]Hwang, C. L., & Lin, M. J. (1987). Group decision making under multiple criteria: Methods and applications. Springer-Verlag.

Iarossi, F. J. (2003). Marine Safety: Perception and Reality. 17th Annual Chua Chor Teck Memorial Lecture. Singapore. Retrieved May 12, 2003 from http://www.eagle.org/news /speeches/index.html.

[13]IMO Model Course on Marine Accident Investigation (2007)

[14]Iranian Ports and Maritime Organization reports on Marine Accident (2012) from http// www.pmo.ir

[15]Ishikawa, A., Amagasa, M., Shiga, T., Tomizawa, G., Tatsuta, R., & Mieno, H. (1993).

[16]J.J. Buckley, (1985) Ranking alternatives using fuzzy numbers, Fuzzy Sets and Systems 15 (1) 21–31.

[17] JAO-HONG CHENG and et al ,An Application of Fuzzy Delphi and Fuzzy AHP on Evaluating Wafer Supplier in Semiconductor Industry ,Department of Information Management National Yunlin University of Science and Technology ,Taiwan

[18]Jones, M. (2002). Review and Analysis of Accident Incident and Near-Miss Databases.

Kahraman, C., Cebeci, U., & Ruan, D. (2004). Multi-attribute comparison of catering service companies using fuzzy AHP:

The case of Turkey. International Journal Production Economics, 87, 171–184.
[19]Klir, G. J., & Yuan, B. (1995). Fuzzy sets and fuzzy logic – Theory and application. New Jersey: Prentice-Hall Inc.
[20]Kristiansen Svein, (2005), Maritime Transportation Safety Management and Risk Analysis 523, Elsevier Butterworth-Heinemann, UK.
[21]Kuei-Yang Wu , Applying the Fuzzy Delphi Method to Analyze the Evaluation In for Service Quality after Railway Re-Opening Using the O Mountain Line Railway as an Example Department of Architecture National United University No. 1 Lien-Da, Kung-Ching Li, Miaoli 36003, Taiwan, ROC Taiwan, ROC kyw@nuu.edu.tw
Laarhoven, P. J. M., & Pedrycz, W. (1983). A fuzzy extension of Sati's priority theory. Fuzzy Sets and System, 11, 229–241.
Lin, B., Lin, C.-Y., & Jong, T.-C. (2007). Investigation of strategies to improve the recycling effectiveness of waste oil from fishing vessels. Marine Policy, 31(4), 415–420.
Linn, Johannes F, "Anticorruption in Transition" (World Bank Publications 2000) ODEC Principals of Corporate Governance
Liu YC, Chen CS (2007), A new approach for application of rock mass classification on rock slope stability assessment, engineering geology 89,pp 129-143
Maritime Accident Review 2010, , European Maritime Safety Agency Maritime Safety Committee (MSC).255(84) Resolution (adopted on 16 May 2008)
Murray, T. J., Pipino, L. L., & Gigch, J. P. (1985). A pilot study of fuzzy set modification of Delphi. Human Systems Management, 6–80.
Reza, K., & Vassilis, S. M. (1988). Delphi hierarchy process (DHP): A methodology for priority setting derived from the Delphi method and analytical hierarchy process. European Journal of Operational Research, 137, 347–354.
Teng, J. Y., & Tzeng, G. H. (1996). Fuzzy Multi Criteria Ranking of urban transportation investment alternative. Transportation Planning and Technology, 20,
The max–min Delphi method and fuzzy Delphi method via fuzzy integration. Fuzzy Sets and Systems,55,241–253.
UK Marine Accident Investigation Branch (MAIB),Annual Report 1999. Retrieved July 16,2003from http://www.maib.detr.gov.uk/ar1999/04.htm
United States Coast Guard Research and Development Center, (2001). United States Coast Guard guide for the management of crew endurance and risk factors (Report No. CG-D-13-01).
What is Good Governance Stephen P. Osborne The New Public Governance, Emerging Perspectives on the Theory and Practice of Public Governance Published In: United Kingdom, 21 December 2009
Zadeh, L. A. (1965). Fuzzy sets. Information Control, 8, 338–353.
Zhau, R., & Goving, R. (1991). Algebraic characteristics of extended fuzzy numbers. Information Science, 54, 103–130.

Authors Index

ISBN: 978-1-61804-227-9