

**IJCSIS Vol. 16 No. 1, January 2018**  
**ISSN 1947-5500**

# **International Journal of Computer Science & Information Security**

**© IJCSIS PUBLICATION 2018**  
**Pennsylvania, USA**

*Indexed and technically co-sponsored by :*



Please consider to contribute to and/or forward to the appropriate groups the following opportunity to submit and publish original scientific results.

## CALL FOR PAPERS

### International Journal of Computer Science and Information Security (IJCSIS) January-December 2018 Issues

The topics suggested by this issue can be discussed in term of concepts, surveys, state of the art, research, standards, implementations, running experiments, applications, and industrial case studies. Authors are invited to submit complete unpublished papers, which are not under review in any other conference or journal in the following, but not limited to, topic areas.

See authors guide for manuscript preparation and submission guidelines.

**Indexed by Google Scholar, DBLP, CiteSeerX, Directory for Open Access Journal (DOAJ), Bielefeld Academic Search Engine (BASE), SCIRUS, Scopus Database, Cornell University Library, ScientificCommons, ProQuest, EBSCO and more.**

**Deadline:** see web site

**Notification:** see web site

**Revision:** see web site

**Publication:** see web site

Context-aware systems

Networking technologies

Security in network, systems, and applications

Evolutionary computation

Industrial systems

Evolutionary computation

Autonomic and autonomous systems

Bio-technologies

Knowledge data systems

Mobile and distance education

Intelligent techniques, logics and systems

Knowledge processing

Information technologies

Internet and web technologies, IoT

Digital information processing

Cognitive science and knowledge

Agent-based systems

Mobility and multimedia systems

Systems performance

Networking and telecommunications

Software development and deployment

Knowledge virtualization

Systems and networks on the chip

Knowledge for global defense

Information Systems [IS]

IPv6 Today - Technology and deployment

Modeling

Software Engineering

Optimization

Complexity

Natural Language Processing

Speech Synthesis

Data Mining

For more topics, please see web site <https://sites.google.com/site/ijcsis/>

arXiv.org Google scholar

SCIRUS  
search engine for science

ScientificCommons

Scribd

.docstoc  
find and share professional documents

BASE  
Bielefeld Academic Search Engine

CiteSeerX beta

uni-trier.de  
Computer Science Bibliography

DOAJ  
DIRECTORY OF OPEN ACCESS JOURNALS

EBSCO HOST

ProQuest

## Editorial Message from Editorial Board

*It is our great pleasure to present the January 2018 issue (Volume 16 Number 1) of the International Journal of Computer Science and Information Security (IJCSIS). High quality research, survey & review articles are proposed from experts in the field, promoting insight and understanding of the state of the art, and trends in computer science and technology. It especially provides a platform for high-caliber academics, practitioners and PhD/Doctoral graduates to publish completed work and latest research outcomes. According to Google Scholar, up to now papers published in IJCSIS have been cited over 9850 times and this journal is experiencing steady and healthy growth. Google statistics shows that IJCSIS has established the first step to be an international and prestigious journal in the field of Computer Science and Information Security. There have been many improvements to the processing of papers; we have also witnessed a significant growth in interest through a higher number of submissions as well as through the breadth and quality of those submissions. IJCSIS is indexed in major academic/scientific databases and important repositories, such as: Google Scholar, Thomson Reuters, ArXiv, CiteSeerX, Cornell's University Library, Ei Compendex, ISI Scopus, DBLP, DOAJ, ProQuest, ResearchGate, LinkedIn, Academia.edu and EBSCO among others.*

*A great journal cannot be made great without a dedicated editorial team of editors and reviewers. On behalf of IJCSIS community and the sponsors, we congratulate the authors and thank the reviewers for their outstanding efforts to review and recommend high quality papers for publication. In particular, we would like to thank the international academia and researchers for continued support by citing papers published in IJCSIS. Without their sustained and unselfish commitments, IJCSIS would not have achieved its current premier status, making sure we deliver high-quality content to our readers in a timely fashion.*

*"We support researchers to succeed by providing high visibility & impact value, prestige and excellence in research publication." We would like to thank you, the authors and readers, the content providers and consumers, who have made this journal the best possible.*

*For further questions or other suggestions please do not hesitate to contact us at [ijcsiseditor@gmail.com](mailto:ijcsiseditor@gmail.com).*

*A complete list of journals can be found at:*

*<http://sites.google.com/site/ijcsis/>*

*IJCSIS Vol. 16, No. 1, January 2018 Edition*

**ISSN 1947-5500 © IJCSIS, USA.**

*Journal Indexed by (among others):*



**Open Access** This Journal is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source.



**Bibliographic Information**

ISSN: 1947-5500

Monthly publication (Regular Special Issues)  
Commenced Publication since May 2009

**Editorial / Paper Submissions:**

**IJCSIS Managing Editor**

[\(ijcsiseditor@gmail.com\)](mailto:ijcsiseditor@gmail.com)

**Pennsylvania, USA**

**Tel: +1 412 390 5159**

# IJCSIS EDITORIAL BOARD

IJCSIS Editorial Board	IJCSIS Guest Editors / Associate Editors
<b>Dr. Shimon K. Modi</b> <a href="#">[Profile]</a> Director of Research BSPA Labs, Purdue University, USA	<b>Dr Riktesh Srivastava</b> <a href="#">[Profile]</a> Associate Professor, Information Systems, Skyline University College, Sharjah, PO 1797, UAE
<b>Professor Ying Yang, PhD.</b> <a href="#">[Profile]</a> Computer Science Department, Yale University, USA	<b>Dr. Jianguo Ding</b> <a href="#">[Profile]</a> Norwegian University of Science and Technology (NTNU), Norway
<b>Professor Hamid Reza Naji, PhD.</b> <a href="#">[Profile]</a> Department of Computer Engineering, Shahid Beheshti University, Tehran, Iran	<b>Dr. Naseer Alquraishi</b> <a href="#">[Profile]</a> University of Wasit, Iraq
<b>Professor Yong Li, PhD.</b> <a href="#">[Profile]</a> School of Electronic and Information Engineering, Beijing Jiaotong University, P. R. China	<b>Dr. Kai Cong</b> <a href="#">[Profile]</a> Intel Corporation, & Computer Science Department, Portland State University, USA
<b>Professor Mokhtar Beldjehem, PhD.</b> <a href="#">[Profile]</a> Sainte-Anne University, Halifax, NS, Canada	<b>Dr. Omar A. Alzubi</b> <a href="#">[Profile]</a> Al-Balqa Applied University (BAU), Jordan
<b>Professor Yousef Farhaoui, PhD.</b> Department of Computer Science, Moulay Ismail University, Morocco	<b>Dr. Jorge A. Ruiz-Vanoye</b> <a href="#">[Profile]</a> Universidad Autónoma del Estado de Morelos, Mexico
<b>Dr. Alex Pappachen James</b> <a href="#">[Profile]</a> Queensland Micro-nanotechnology center, Griffith University, Australia	<b>Prof. Ning Xu,</b> Wuhan University of Technology, China
<b>Professor Sanjay Jasola</b> <a href="#">[Profile]</a> Gautam Buddha University	<b>Dr . Bilal Alatas</b> <a href="#">[Profile]</a> Department of Software Engineering, Firat University, Turkey
<b>Dr. Siddhivinayak Kulkarni</b> <a href="#">[Profile]</a> University of Ballarat, Ballarat, Victoria, Australia	<b>Dr. Ioannis V. Koskosas,</b> University of Western Macedonia, Greece
<b>Dr. Reza Ebrahimi Atani</b> <a href="#">[Profile]</a> University of Guilan, Iran	<b>Dr Venu Kuthadi</b> <a href="#">[Profile]</a> University of Johannesburg, Johannesburg, RSA
<b>Dr. Dong Zhang</b> <a href="#">[Profile]</a> University of Central Florida, USA	<b>Dr. Zhihan Lv</b> <a href="#">[Profile]</a> Chinese Academy of Science, China
<b>Dr. Vahid Esmaeilzadeh</b> <a href="#">[Profile]</a> Iran University of Science and Technology	<b>Prof. Ghulam Qasim</b> <a href="#">[Profile]</a> University of Engineering and Technology, Peshawar, Pakistan
<b>Dr. Jiliang Zhang</b> <a href="#">[Profile]</a> Northeastern University, China	<b>Prof. Dr. Maqbool Uddin Shaikh</b> <a href="#">[Profile]</a> Preston University, Islamabad, Pakistan
<b>Dr. Jacek M. Czerniak</b> <a href="#">[Profile]</a> Casimir the Great University in Bydgoszcz, Poland	<b>Dr. Musa Peker</b> <a href="#">[Profile]</a> Faculty of Technology, Mugla Sitki Kocman University, Turkey
<b>Dr. Binh P. Nguyen</b> <a href="#">[Profile]</a> National University of Singapore	<b>Dr. Wencan Luo</b> <a href="#">[Profile]</a> University of Pittsburgh, US
<b>Professor Seifeidne Kadry</b> <a href="#">[Profile]</a> American University of the Middle East, Kuwait	<b>Dr. Ijaz Ali Shoukat</b> <a href="#">[Profile]</a> King Saud University, Saudi Arabia
<b>Dr. Riccardo Colella</b> <a href="#">[Profile]</a> University of Salento, Italy	<b>Dr. Yilun Shang</b> <a href="#">[Profile]</a> Tongji University, Shanghai, China
<b>Dr. Sedat Akylek</b> <a href="#">[Profile]</a> Ondokuz Mayis University, Turkey	<b>Dr. Sachin Kumar</b> <a href="#">[Profile]</a> Indian Institute of Technology (IIT) Roorkee

<b>Dr Basit Shahzad</b> [ <a href="#">Profile</a> ] King Saud University, Riyadh - Saudi Arabia	<b>Dr. Mohd. Muntjir</b> [ <a href="#">Profile</a> ] Taif University Kingdom of Saudi Arabia
<b>Dr. Sherzod Turaev</b> [ <a href="#">Profile</a> ] International Islamic University Malaysia	<b>Dr. Bohui Wang</b> [ <a href="#">Profile</a> ] School of Aerospace Science and Technology, Xidian University, P. R. China
<b>Dr. Kelvin LO M. F.</b> [ <a href="#">Profile</a> ] The Hong Kong Polytechnic University, Hong Kong	

ISSN 1947 5500 Copyright © IJCSIS, USA.

## TABLE OF CONTENTS

### 1. PaperID 31121708: Coding Instead of Encryption (pp. 1-7)

*Dr. Yaseen Hikmat Ismaiel,  
Department Computer Science, College of Computer and Mathematic Science, Mosul University*

Full Text: PDF [Academia.edu | Scopus | Scribd | Archive | ProQuest]

### 2. PaperID 31121712: Three Dimensional Chaotic System for Color Image Scrambling Algorithm (pp. 8-20)

*S. N. Lagmiri (1), N. Elalami (2), J. Elalami (3)  
(1) Information and Production System, Mohammadia School Engineering, Mohamed V University in Rabat, Morocco  
(2) LAII, Mohammadia School Engineering, Mohamed V University in Rabat, Morocco  
(3) LASTIMI, Higher School of Technology of Sale, Mohamed V University in Rabat, Morocco*

Full Text: PDF [Academia.edu | Scopus | Scribd | Archive | ProQuest]

### 3. PaperID 31121719: Data Partitioning In Cloud Storage Using DESD Crypto Technique (pp. 21-30)

*B. Muthulakshmi (1)\*, M. Venkatesulu (2)  
(1)\* Department of Computer Applications, Kalasalingam University, Krishnankoil, Tamil Nadu, India.  
(2) Department of Information Technology, Kalasalingam University, Krishnankoil, Tamil Nadu, India.*

Full Text: PDF [Academia.edu | Scopus | Scribd | Archive | ProQuest]

### 4. PaperID 31121721: Design a Hybrid Technique Based new Genetic Approach for Text Encryption (pp. 31-37)

*Shatha Abdullah M. Ramadhan, Dept. of Software Engineering, College of Computer Sc. & Math., University of Mosul, Mosul, Iraq.*

*Hiba Muneer Yahya, Dept. of Software Engineering, College of Computer Sc. & Math., University of Mosul, Mosul, Iraq.*

*Ruah Mouad Alyas, College of Computer Sc. & Math., University of Mosul, Mosul, Iraq*

Full Text: PDF [Academia.edu | Scopus | Scribd | Archive | ProQuest]

### 5. PaperID 31121723: Analyzing the Impacts of Solar PV Generation Integration on Radial Distribution Feeder (pp. 38-42)

*Mohsin A. Mari, Department of Electrical Engineering, Mehran UET SZAB Campus Khairpur Mir's  
Anwar A. Sahito, Department of Electrical Engineering, Mehran UET, Jamshoro*

*M. Zubair Bhayo, Department of Electrical Technology, Benazir Bhutto Shaheed University of Technology and Skill Development Khairpur Mir's*

*Basheer A. Kalwar, Department of Electrical Engineering, Mehran UET SZAB Campus Khairpur Mir's*

Full Text: PDF [Academia.edu | Scopus | Scribd | Archive | ProQuest]

**6. PaperID 31121725: Which NoSQL Database to Combine with Spark for Real Time Big Data Analytics? (pp. 43-46)**

*Omar HAJOUI & Mohamed TALEA,  
LTI Laboratory, Faculty of Science Ben M'Sik, Hassan II University, Casablanca, Morocco*

[Full Text: PDF](#) [[Academia.edu](#) | [Scopus](#) | [Scribd](#) | [Archive](#) | [ProQuest](#)]

**7. PaperID 31121729: Scheduling Problems from Workshop to Collaborative Mobile Computing: A State of the Art (pp. 47-69)**

*Ghassan Fadlallah, Hamid Mccheick, Djamel Rebaine  
Département d'informatique et de mathématique, University of Quebec at Chicoutimi 555, boul. De l'Université,  
Saguenay, Québec, Canada G7H 2B1*

[Full Text: PDF](#) [[Academia.edu](#) | [Scopus](#) | [Scribd](#) | [Archive](#) | [ProQuest](#)]

**8. PaperID 31121730: Digitalization and Education System: A Survey (pp. 70-73)**

*Saiful Islam, Lecturer at dept. of CSE, Daffodil International University, Dhaka, Bangladesh  
Nusrat Jahan, Dept. of CSE, Daffodil International University, Dhaka, Bangladesh*

[Full Text: PDF](#) [[Academia.edu](#) | [Scopus](#) | [Scribd](#) | [Archive](#) | [ProQuest](#)]

**9. PaperID 31121731: Implementation on Data Security Approach in Dynamic Multi Hop Communication (pp. 74-79)**

*Dr. S. A. Ubale, Zeal college of Engineering and research, Narhe 41, Pune, India  
Anuradha V. Khare, Zeal college of Engineering and research, Narhe 41, Pune, India  
Dr. Sulbha S. Apte, Walchand Institute of Technology, Solapur 06, Solapur, India*

[Full Text: PDF](#) [[Academia.edu](#) | [Scopus](#) | [Scribd](#) | [Archive](#) | [ProQuest](#)]

**10. PaperID 31121733: LCF: A Temporal Approach to Link Prediction in Dynamic Social Networks (pp. 80-88)**

*A. Jagadishwari. V, Research Scholar, VTU, BMS College of Engineering, Bengaluru  
B. Umadevi. V, BMS College of Engineering, Bengaluru*

[Full Text: PDF](#) [[Academia.edu](#) | [Scopus](#) | [Scribd](#) | [Archive](#) | [ProQuest](#)]

**11. PaperID 31121734: An Implementation of a New Framework for Automatic Generation of Ontology and RDF to Real-Time Web and Journal Data (pp. 89-94)**

*Sudeepthi Govathoti (1), M.S. Prasad Babu (2)  
(1) Research Scholar, Andhra University, Visakhapatnam &Assistant Professor, Anurag Group of Institutions,  
Hyderabad,  
(2) Professor, Department of CS&SE, Andhra University, Visakhapatnam,*

[Full Text: PDF](#) [[Academia.edu](#) | [Scopus](#) | [Scribd](#) | [Archive](#) | [ProQuest](#)]

**12. PaperID 31121737: Nonnegative Garrote as a Variable Selection Method in Panel Data (pp. 95-106)**

*Borislava Vrigazova, Sofia University “St. Kliment Ohridski”*

**Full Text:** PDF [Academia.edu | Scopus | Scribd | Archive | ProQuest]

**13. PaperID 31121738: Implementing a Solution to the Cloud Vendor Lock-In Using Standardized API (pp. 107-111)**

*K. V. V. N. Prasanth, K. V. V. Satyanarayana, V. Akhila, M. Sahithya, A. Sandeep Reddy  
Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram-522502, Guntur Dt., AP*

**Full Text:** PDF [Academia.edu | Scopus | Scribd | Archive | ProQuest]

**14. PaperID 31121739: Non-Separable Histogram Based Reversible Data Hiding Approach Using Inverse S-order and Skew Tent Map (pp. 112-127)**

*Anoop Kumar Chaturvedi, Research Scholar, UIT, RGPV, Bhopal  
Dr. Piyush Kumar Shukla, DoCSE,UIT, RGPV, Bhopal*

**Full Text:** PDF [Academia.edu | Scopus | Scribd | Archive | ProQuest]

**15. PaperID 31121741: Fractal Compression of an AVI Video File using DWT and Particle Swarm Optimization (pp. 128-131)**

*Shraddha Pandit, Research Scholar, Department of Computer Science & Engineering, University Institute of Technology, Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal, India  
Piyush Kumar Shukla, Assistant Professor, Department of Computer Science & Engineering, University Institute of Technology, Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal, India  
Akhilesh Tiwari, Associate Professor, Department of Computer Science & Engineering and Information Technology, Madhav Institute of Technology & Science, Gwalior, India*

**Full Text:** PDF [Academia.edu | Scopus | Scribd | Archive | ProQuest]

**16. PaperID 31121744: A Proposal on a Method of ArchiMate based Concept of Operation (ConOps) (pp. 132-138)**

*Nada Olayan, Graduate School of information science, Nagoya University, 464-8601 Nagoya, Japan  
Shuji Morisaki, Graduate School of information science, Nagoya University, 464-8601 Nagoya, Japan  
Nobuhide Kobayashi, Denso Create Inc., 460-0008 Nagoya, Japan  
Shuichiro Yamamoto, Graduate School of information science, Nagoya University, 464-8601 Nagoya, Japan*

**Full Text:** PDF [Academia.edu | Scopus | Scribd | Archive | ProQuest]

**17. PaperID 31121750: Security Level Analysis of Academic Information Systems Based on Standard ISO 27002: 2013 using SSE-CMM (pp. 139-147)**

*Endang Kurniawan, Department of Informatics Engineering, University Islam of Indonesia, Jogjakarta, Indonesia  
Imam Riadi, Department of Information System, Ahmad Dahlan University, Jogjakarta, Indonesia*

[Full Text: PDF](#) [[Academia.edu](#) | [Scopus](#) | [Scribd](#) | [Archive](#) | [ProQuest](#)]

**18. PaperID 31121751: Simulation and Coding of a Neural Network, Performing Generalized Function with Real Inputs (pp. 148-157)**

*Elitsa Zdravkova Spasova, Faculty of Mathematics and Informatics, University of Shumen “Episkop Konstantin Preslavsky”, 9700, Shumen, Bulgaria*

[Full Text: PDF](#) [[Academia.edu](#) | [Scopus](#) | [Scribd](#) | [Archive](#) | [ProQuest](#)]

**19. PaperID 31121752: Using Genetic Algorithm for Shortest Path Selection with Real Time Traffic Flow (pp. 158-162)**

*Mohammad Aldabbagh, College of Computer Science and Mathematics, University of Mosul*

[Full Text: PDF](#) [[Academia.edu](#) | [Scopus](#) | [Scribd](#) | [Archive](#) | [ProQuest](#)]

**20. PaperID 31121753: A Software Infrastructure for Multidimensional Data Analysis: A Data Modelling Aspect (pp. 163-171)**

*Prarthana A. Deshkar, Assistant Professor Yeshwantrao Chavan College of Engineering Nagpur, India  
Dr. Parag S. Deshpande, Supervisor, CSE dept, G. H. Raisoni College of Engineering, Nagpur, India  
Prof. A. Thomas, HoD, CSE department G. H. Raisoni College of Engineering, Nagpur, India*

[Full Text: PDF](#) [[Academia.edu](#) | [Scopus](#) | [Scribd](#) | [Archive](#) | [ProQuest](#)]

**21. PaperID 31121749: Enhanced Generalized Regression Neural Net for Breast Cancer Detection (pp. 172-173)**

*Samira Babaei Ghalejoughi, Department of Electronic Engineering, Tabriz Branch, Islamic Azad University, Tabriz, Iran.  
Nasser Lotfivand, Department of Electronic Engineering, Tabriz Branch, Islamic Azad University, Tabriz, Iran*

[Full Text: PDF](#) [[Academia.edu](#) | [Scopus](#) | [Scribd](#) | [Archive](#) | [ProQuest](#)]

# Coding instead of encryption

Dr. Yaseen Hikmat Ismaiel

Lecture

Department computer science

Collage of computer and mathematic science

Mosul University

[Yasino79@yahoo.com](mailto:Yasino79@yahoo.com)

## ABSTRACT

Due to the increased use of the internet in recent decades and the large number of transactions and data exchanged over the network . There is an urgent need to provide security for those data, especially those related to military, commercial, or financial exchanges. Many cryptographic methods have recently been widely investigated and developed because there is a demand for a stronger encryption and decryption which is very hard to crack. Most modern encrypting methods include many substitution and iteration processes that encountered some problems such as lack of robustness and significant amount of time.

In this research, a proposed method based on the idea of coding was introduced to achieve security instead of encryption. This search uses ASCII to build a coding table in a different way to provide security and saves effort, time and cost.

**Keywords :** coding theory , code book , encryption.

## 1. Introduction :

It is well known that cryptographic systems use a key and algorithm to convert plain text into ambiguous (cipher) text, while encoding methods include converting plain text to another format using a particular method or special coding table. The purpose of coding systems may not be to achieve security but to convert data to another format and use it or take advantage of them in a particular area. Examples of coding methods include ASCII , Unicode , URL encoding , base64 [1] [2] [3] [4].

This paragraph includes a review of the studies and research in this field, where the researcher in 1978 McEliece [5] using the fact that a fast decoding algorithm exists for general Goppa code, he constructed a public-key cryptosystem using algebraic coding theory . In 2005 Grangetto M. and et al [6] introduced a randomized arithmetic coding paradigm, which achieves encryption by inserting some randomization in the arithmetic coding procedure , the proposed approach allows very flexible protection procedures at the code-block level, allowing to perform total and selective encryption, as well as conditional access. In 2010 Wong K. and et al [7] proposed a simultaneous compression and encryption scheme in which the chaotic map model for arithmetic coding is determined by a secret key and keeps changing. In 2010 Lu R. and et al [8] proposed an efficient INDCCA2-secure public key encryption scheme based on coding theory, and then measure the efficiency of the method by comparing with the syndrome decoding problem in the random oracle model. In 2012 Gupta V. and et al [9] proposed an improved block cipher symmetric encryption algorithm that has the same structure of encryption and decryption by inserting the symmetric layer; The method gave high efficiency in term of speed. In 2013 Al-Hazaimah O. [10] presented a new approach based on parallel programming to

provide complexity in encryption and decryption process ; it offers a high level of security with good speed .

## **2. Coding and Encryption System :**

Coding systems originated long ago and require a mathematical background for the purpose of building a good coding theory. Coding system rely on the use of the dictionary (table) without the need to use a complex key or algorithm, its purpose is to convert data to another format so that it can be used in many applications. [3] [11]

The objective is different from using coding systems; some applications are designed to achieve confidentiality, some to data compression, or to convert data into a specific format that can be used in specific applications. Coding system divided into two types one-part and two-part codes, in the one part type the code groups and the vocabulary are arranged in parallel, alphabetic sequences, so that a single book serves for encoding as well as for decoding. In the two-part type the encoding book lists the elements of the vocabulary in alphabetic order but the code groups are in random order, so that a decoding book, in which the code groups appear in alphabetic (or numerical) order accompanied by their meaning, is essential. The degree of secrecy afforded by a code of the latter type is much greater than that afforded by one of the former type, all other things being equal. [3] [4][11]

Encryption is the process of encoding a message or information in such a way that only authorized parties can access it and those who are not authorized cannot. Encryption is one of the most important methods for providing data privacy(confidentiality) , authenticity (it came from where it claims) , and data integrity(it has not been modified on the way) in the digital world , especially for end-to-end protection of data transmitted across networks. There are many types of encryption systems and certain measures have been used to divide them, it is divided in terms of the nature of its work into transposition and substitution encryption systems, or divided in terms of how to use the key to secret and public encryption systems. [1] [2] [12] [13]

## **3. Proposed Method :**

The proposed method involves using the concept of coding as a good alternative to encryption methods in terms of achieving high level of security and cost savings. It is known that the coding methods used to achieve confidentiality depend on the building of the coding book, which serves as a key to the coding. The mechanism of these methods is to take a word after a word of the plain text and search it in the coding book and then get the word encoded thus configuring the encoded text. The proposed method used the encoding process in a different way depending on ASCII encoding (ASCII printable characters) where the coding table (Table 1) was constructed by using the following rules:

- ✓ Two codes were assigned to the most frequently used characters (E, A, I, N, O, S, and T) to make the resulting symbols in close frequency ratios and thus eliminate the frequency analysis.
- ✓ Codes for commonly used pronouns are assigned (I, me, my, you, your, he, him, she, her, it, its, we, us, our, they, them, their).
- ✓ Encoding question tools (who, what, where, when, why, how, and which).
- ✓ Encoding numbers (0-9).
- ✓ Two codes were assigned to the space also to eliminate the frequency analysis.
- ✓ Codes for commonly used suffixes and prefixes are assigned (able, fully, sion, tion, er, ing, ed, pre, un, re).

- ✓ Encoding week days (Sunday, Monday, Tuesday, Wednesday, Thursday, Friday, and Saturday ).
- ✓ Encoding commonly used prepositions (in, on, and at).
- ✓ Encoding commonly used helping verbs (is, are, has, have, do, does).

### **3.1. The Proposed Method Algorithm :**

The algorithm of the proposed method can be summarized by the following points:

1. Read the text to be encoded.
2. Calculate the number of words in clear text based on the existence of spaces.
3. Search for pronouns, numbers, spaces, weekdays, prepositions, and helping verbs in the clear text and compensate them with the corresponding symbols using table 1 (note the use of the symbols allocated to the space in a sequential manner).
4. Encode the suffixes and prefixes wherever they are in the remaining words and then encode the characters of those remaining words, taking into account when encoding the most frequent characters (E, A, I, N, O, S, and T) using the two codes assigned to them in succession.
5. After finishing the encoding process, the number that represents the number of words in the text is clearly (step 2) encoded and added to the end of the resulting coded text. This number will be used by the recipient to verify the integrity of the received text.

The mechanism of the proposed algorithm can be illustrated in the flowchart in figure 1.

In the receiver, the decoding algorithm can be explained to obtain clear text through the following points:

1. Get the codes at the end of the encoded text and decode them to extract the number that represents the number of clear text words and keep it.
2. Use table 1 to decode all words, sections and characters to get clear text.
3. Calculate the number of clear text words generated and compare them with the number obtained in step 1 if the values are equal, indicating that the received text is correct and the decoding process was done correctly. The flowchart algorithm can be illustrated in figure 2.

Table (1) The coding table used in the suggested method

ascii	chacter	plain	ascii	chacter	plain	ascii	chacter	plain
32		A	64	@	Z	96	`	2
33	!	A	65	A	I	97	a	1
34	"	B	66	B	ME	98	b	0
35	#	C	67	C	MY	99	c	SPACE
36	\$	D	68	D	YOU	100	d	SPACE
37	%	E	69	E	YOUR	101	e	ABLE
38	&	E	70	F	HE	102	f	FULLY
39	,	F	71	G	HIM	103	g	SION
40	(	G	72	H	SHE	104	h	TION
41	)	H	73	I	HER	105	i	ER
42	*	I	74	J	IT	106	j	ING
43	+	I	75	K	ITS	107	k	ED
44	,	J	76	L	WE	108	l	PRE
45	-	K	77	M	US	109	m	UN
46	.	L	78	N	OUR	110	n	RE
47	/	M	79	O	THEY	111	o	SUNDAY
48	0	N	80	P	THEM	112	p	MONDAY
49	1	N	81	Q	THEIR	113	q	TUESDAY
50	2	O	82	R	WHO	114	r	WEDNESDAY
51	3	O	83	S	WHAT	115	s	THURSDAY
52	4	P	84	T	WHERE	116	t	FRIDAY
53	5	Q	85	U	WHEN	117	u	SATURDAY
54	6	R	86	V	WHY	118	v	IN
55	7	S	87	W	HOW	119	w	ON
56	8	S	88	X	WHICH	120	x	AT
57	9	T	89	Y	9	121	y	IS
58	:	T	90	Z	8	122	z	ARE
59	;	U	91	[	7	123	{	HAS
60	<	V	92	¥	6	124		HAVE
61	=	W	93	]	5	125	}	DO
62	>	X	94	^	4	126	~	DOES
63	?	Y	95	_	3			

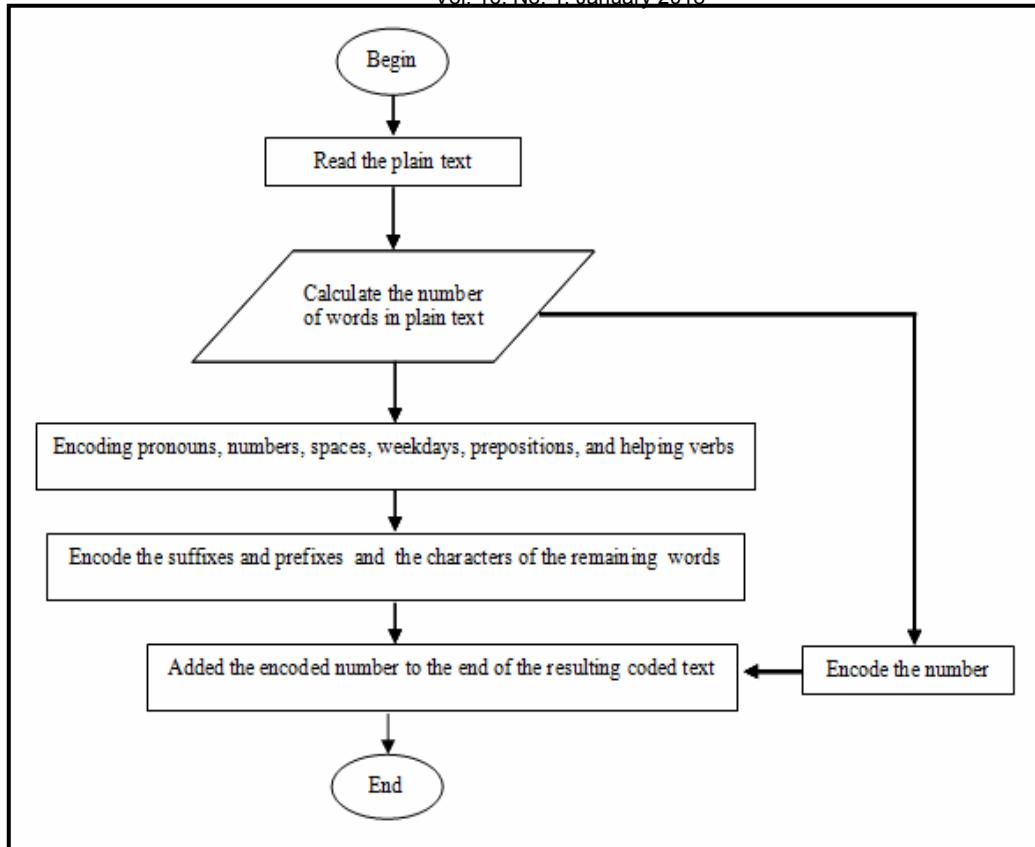


Figure 1 proposed algorithm encoding

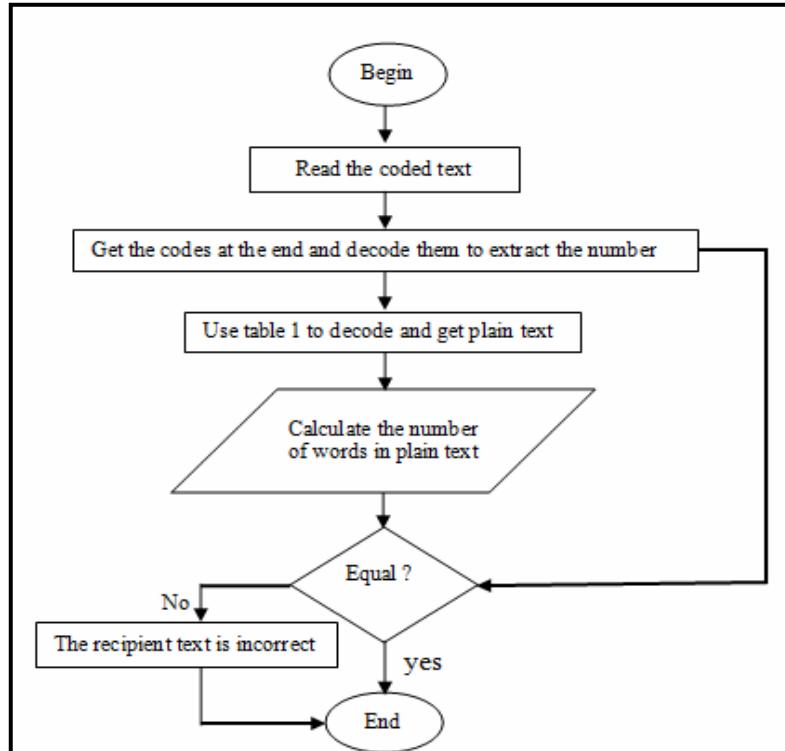


Figure 2 proposed algorithm decoding

#### 4. Conclusion and further work :

- ✓ The use of the coding book in a different way so that some of the words and sections (commonly used in English language) are encoded and then encoding the characters of words that are not in the coding book by encoding (suffixes and prefixes) and then the rest of the word's characters.
- ✓ Calculate the number of words in clear text based on the existence of spaces and inserting the encoded number to the end of the coded text, It provided an innovative method for verifying integration at the receiving party.
- ✓ Assigned two codes to the most frequently used characters (E, A, I, N, O, S, and T) and spaces to make the resulting symbols in close frequency ratios and thus eliminate the frequency analysis.
- ✓ The proposed method provided confidentiality and saves effort, time and cost.
- ✓ The size of the coding book that was built in the proposed method is relatively small "and provided a high level of confidentiality for the resulting coded text.
- ✓ The length of the resulting encoded text is considered to be significantly lower than the length of the plain text, especially when the text is relatively long, so the proposed method provided a good compression ratio in addition to confidentiality.
- ✓ In the future, it is possible to use intelligent techniques in constructing a code book, or to use it for the purpose of converting clear text into another form with the same sense so that the largest number of words have symbols in a book or coding table.

#### 5. References :

- [1] C. Margare and M. Steven, 2013, “The Mathematics of Encryption An Elementary Introduction”, the American Mathematical Society. Mathematical World Volume 29 .
- [2] PGP Corporation , 2002, “An Introduction to Cryptography”, PGP Corporation version 8.0..
- [3] C. julius , 2002, “Codes and ciphers” , the press syndicate of the university of Cambridge.
- [4] F. WILLIAM , 1961, “CODES AND CIPHERS”, National Security Agency  
REF ID:A101014 .
- [5] R. McEliece , 1978, “A Public-key Cryptosystem Based On Algebraic Coding Theory”, DSN Progress .
- [6] G. Marco and et al.,2005, “Multimedia selective encryption by means of randomized arithmetic coding”, IEEE TRANSACTIONS ON MULTIMEDIA .
- [7] W. Kwok and L. Qiuzhen, 2010, “Simultaneous Arithmetic Coding and Encryption Using Chaotic Maps”, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—II: EXPRESS BRIEFS, VOL. 57, NO. 2 .
- [8] L. Rongxing and et al., 2010, “An efficient and provably secure public key encryption scheme based on coding theory”, John Wiley & Sons, Ltd. .
- [9] G. Vishwa and et al., 2012, “A Hyper Modern Cryptography Algorithm to Improved Data Security: HMCA”, International Journal of Computer Science & Communication Networks, Vol 1(3), 258-263 .

- [10] A. Obaida, 2013, “A NEW APPROACH FOR COMPLEX ENCRYPTING AND DECRYPTING DATA”, International Journal of Computer Networks & Communications (IJCNC) Vol.5, No.2 .
- [11] J.H. Ewing and et al. , 1992, “Introduction to Coding Theory”, Springer-Verlag Berlin Heidelberg.
- [12] E. dorothy and D. robling , 1982, “cryptography and data security”, Addison-wesley publishing company, Inc. .
- [13] S. William , 2011, “CRYPTOGRAPHY AND NETWORK SECURITY” , Prentice Hall .

# Three Dimensional Chaotic System for Color Image Scrambling Algorithm

S. N. Lagmiri<sup>1</sup>, N. Elalami<sup>2</sup>, J. Elalami<sup>3</sup>

<sup>1</sup> Information and Production System, Mohammadia School Engineering, Mohamed V University in Rabat, Morocco

<sup>2</sup> LAII, Mohammadia School Engineering, Mohamed V University in Rabat, Morocco

<sup>3</sup> LASTIMI, Higher School of Technology of Sale, Mohamed V University in Rabat, Morocco

*Abstract –With the development of information security, the traditional image encryption methods have become outdated. Because of amply using images in the transmission process, it is important to protect the confidential image data from unauthorized access. This paper presents a new chaos based image encryption algorithm, which can improve the security during transmission more effectively utilizes the chaotic systems properties, such as pseudo-random appearance and sensitivity to initial conditions. Based on chaotic theory and decomposition and recombination of pixel values, this new image scrambling algorithm is able to change the position of pixel, simultaneously scrambling both position and pixel values. Experimental results show that the new algorithm improves the image security effectively to avoid unscramble, and it also can restore the image as same as the original one, which reaches to the purposes of image safe and reliable transmission.*

**Keywords:** Color image, chaotic system, decomposition, image scrambling, recombination

## I. INTRODUCTION

Recently, security of multimedia data is receiving more and more attention due to the transmission over various communication networks. In order to protect personal information, many image encryption algorithms are designed and proposed such as two-dimensional cellular automata based method [2], Henon chaotic map [10, 13], Chen's hyperchaotic system [12], Arnold transform [3, 4]. Chaotic functions are blessed with properties like sensitivity to the initial conditions, and ergodicity which make them very desirable for encryption [1].

Image scrambling is one of the methods for securing the image by scrambling it into a disordered one beyond recognition, making it hard for those who get the image in unauthorized manner to extract information of the original image from the scrambled images. Further, image scrambling technology depends on data hiding technology which provides non-password security algorithm for information hiding. Now, the mainly used three kind of image scrambling types are scrambling in the space domain, scrambling in the frequency domain, and scrambling in the color or grey domain. In a great quantity of all kind of image scrambling algorithms, the image scrambling algorithms based on chaos have attracted more and more attention since they can provide a high level of security [5, 6, 7, 8].

This paper focuses on a new image scrambling algorithm which introduces a new chaotic system. Image scrambling using chaotic properties is an application for providing security to the images from getting into the hands of unauthorized user. The proposed image scrambling scheme generates the permuting address codes by sorting the chaotic sequence directly. This paper analyzed that the scrambling performance of the new algorithm is statistic. The conclusion of this paper indicates that the new algorithm can provide a high level security. The paper results in good performance of the proposed algorithm that can also be applied in the real-time applications and digital communications as it is a straightforward mechanism and easy to implement.

The rest of the paper is organized as follows: proposed chaotic system in section 2, image scrambling algorithm based on chaos theory in section 3, experimental details and results are analyzed in section 4. The paper is observed by a conclusion in section 5.

## II. PROPOSED CHAOTIC SYSTEM

In this section, we describe the new chaotic system used in this work.

### 2.1. New Chaotic System

Recently, Chen and Lee [9] introduced a new chaotic system, which is described by the following nonlinear differential equation:

$$\begin{cases} \dot{x}_1 = ax_1 - x_2x_3 \\ \dot{x}_2 = -bx_2 + x_1x_3 \\ \dot{x}_3 = -cx_3 + \frac{1}{3}x_1x_2 \end{cases} \quad (1)$$

$$W = Cx$$

Where:

-  $x_1$ ,  $x_2$  and  $x_3$  are the state variables and  $a$ ,  $b$  and  $c$  are positive constants.

-  $C = (1 \ 0 \ 0)$ .

-  $W$  is the system measured output.

When  $a = 5.5$ ,  $b = 11$  and  $c = 4$ , the system (1) is chaotic.

### 2.2. Lyapunov exponent

By linearizing the Jacobian matrix  $J_E$  round the equilibrium point  $E$  and solving the following equation:

$$|\lambda_i - J_E| = 0 \quad (2)$$

Therefore, the new chaotic system (1) has three eigenvalues shown in figure 1.

$$\lambda_1 = 5.500 \quad \lambda_2 = -10.994 \quad \lambda_3 = -3.999$$

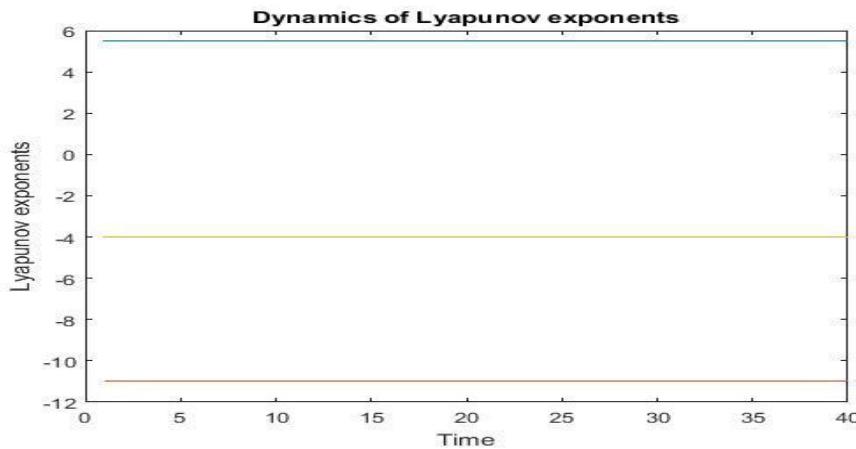
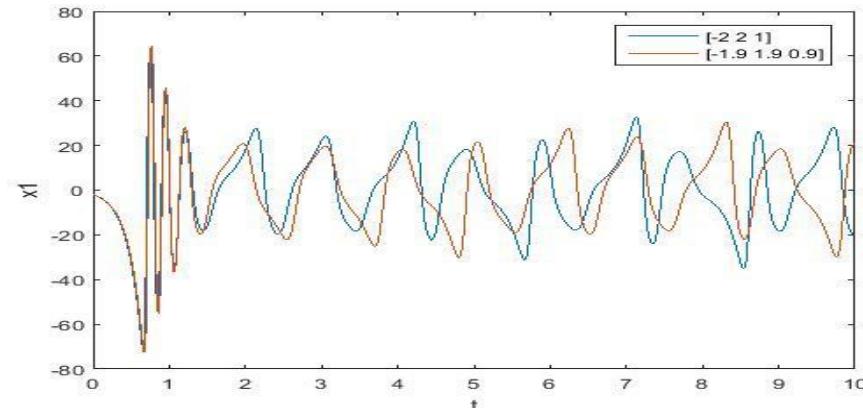


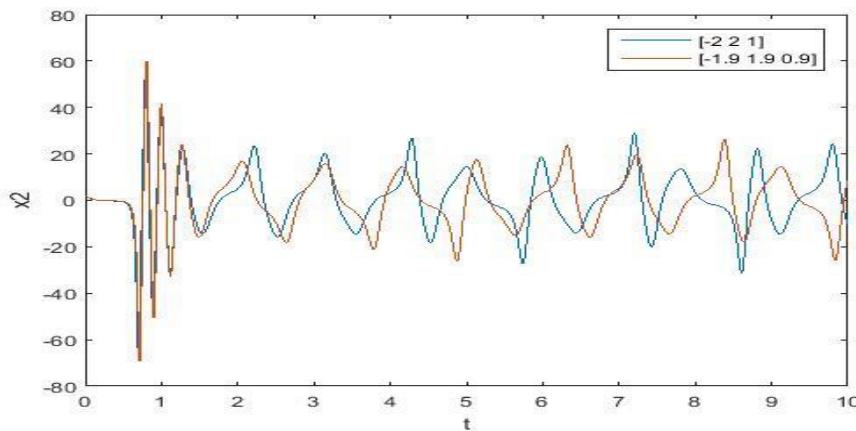
Figure 1: Lyapunov exponent of new chaotic system

### 2.3. Sensitivity to initial conditions

Sensitivity to initial conditions means that each point in a chaotic system is arbitrarily closely approximated by other points with significantly different future paths, or trajectories. Thus, an arbitrarily small change, or perturbation, of the current trajectory may lead to significantly different future behavior. The next figure compares the time series for two likely different initial conditions. The two time series stay close together for about 2 iterations. But after that, they are pretty much on their own.



(a)



(b)

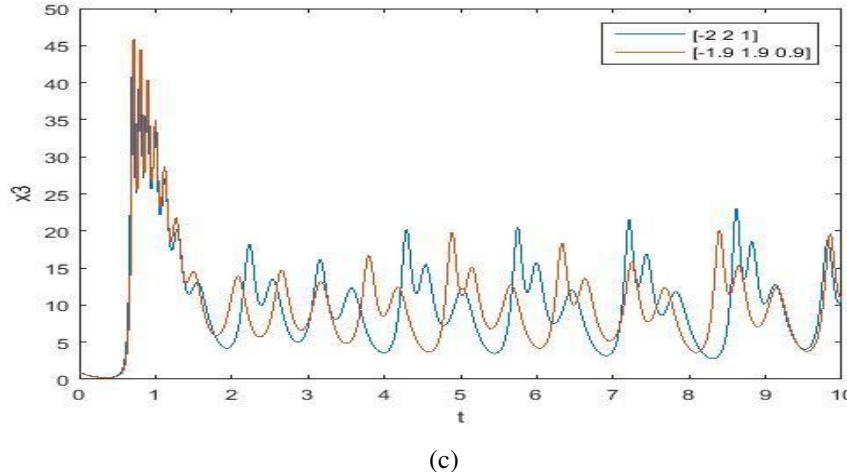


Figure 2: Sensitivity to two initial conditions [-2, 2, 1] and [-1.9, 1.9, 0.9]

(a):  $x_1$  (b):  $x_2$  (c):  $x_3$

### III. IMAGE SCRAMBLING ALGORITHM BASED ON CHAOS THEORY

#### 3.1. Proposed algorithm

The proposed chaotic system is now used in the design of color image encryption algorithm. The proposed images encryption algorithm input is an original image whilst the output is a scrambled one. Figure 3 illustrate the proposed algorithm scheme.

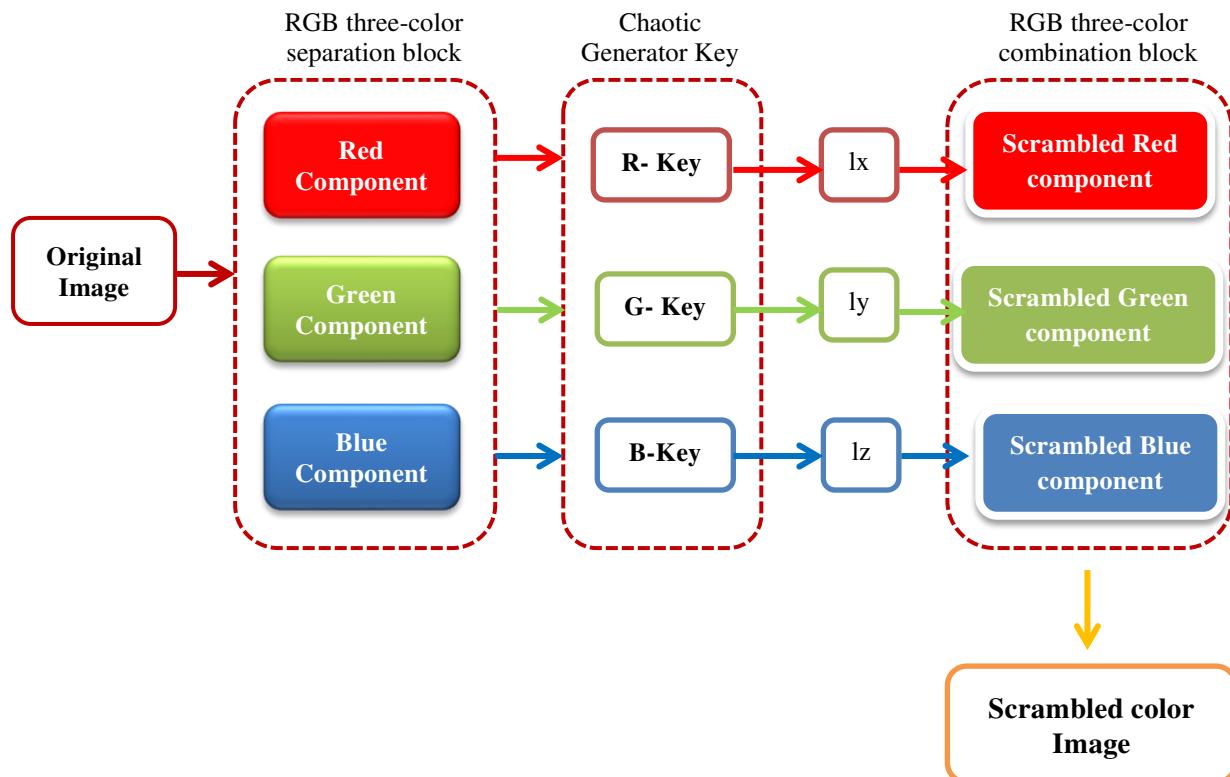


Figure 3: Principle of chaotic scrambling algorithm for color image

#### IV. EXPERIMENTAL DETAILS AND RESULTS

A good quality scrambled algorithm should be strong against all types of attack. Some experiments are given in this section to demonstrate the efficiency of the proposed technique. In this section, the proposed technique is applied on two color images "Gallery" and "Alice", of resolution of " 256\*256". We analyze the results by calculating histogram and correlation coefficient, to test the performance of the proposed technique. The next figures show the results of scrambled algorithm.

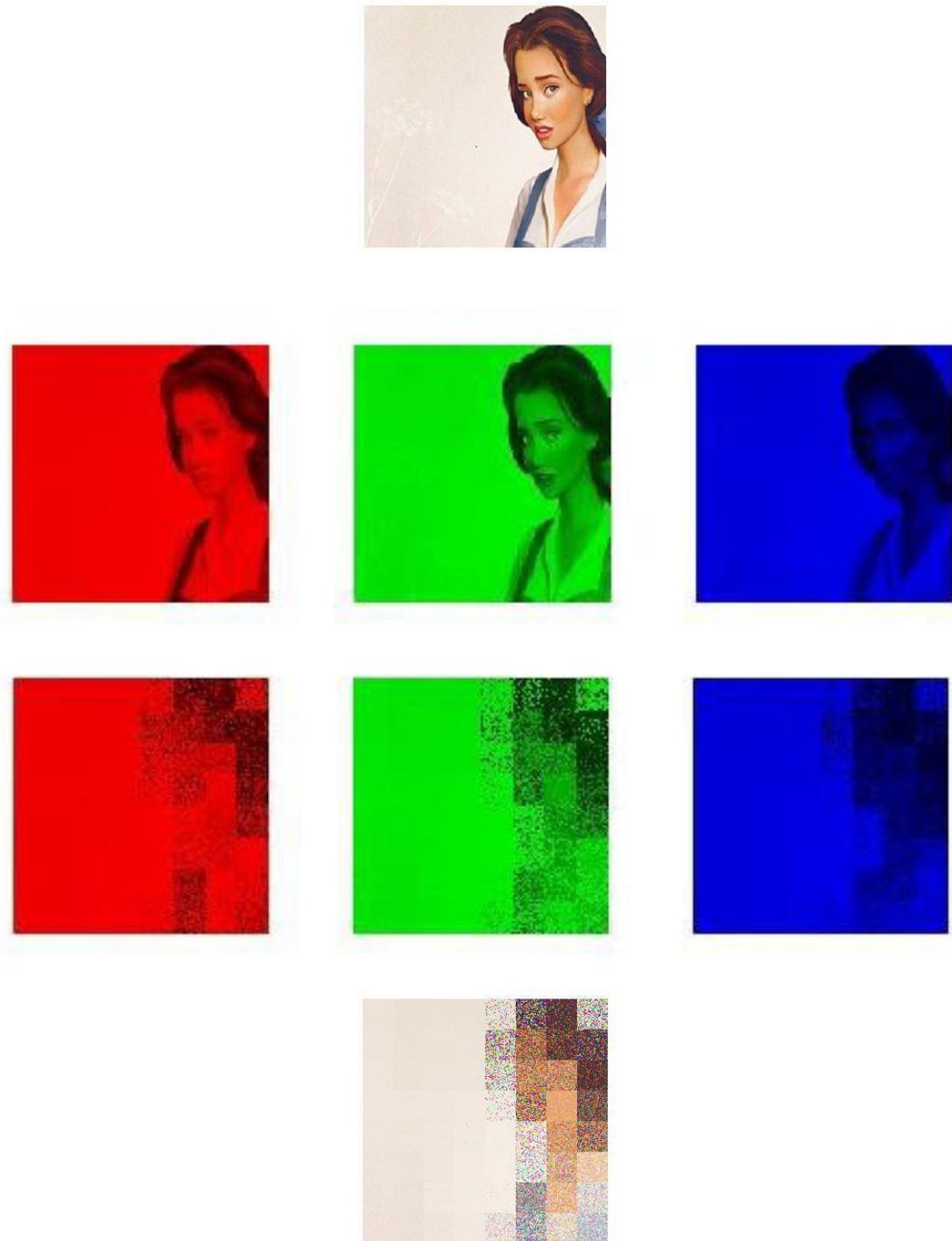


Figure 4: "Alice" image corresponding for different step of the scrambling process



Figure 5: "Gallery" image corresponding for different step of the scrambling process

#### 4.1. Statistical analysis

In order to resist attacks, the scrambled images should possess certain random properties. To prove the robustness of the proposed algorithm, a statistical analysis has been performed by calculating the histograms and the correlation coefficients for the original image and the scrambled image. For the two images that have been tested, it has been determined that their quality is good.

#### 4.1.1. Histogram Analysis

An image histogram is a commonly used method of analysis in image processing. The advantage of a histogram is that it shows the shape of the distribution for a large set of data. Thus, an image histogram illustrates how pixels in an image are distributed by plotting the number of pixels at each color intensity level. It is important to ensure that the encrypted and original images do not have any statistical similarities.

The experimental results of the original image and its corresponding scrambled image and their histograms are shown in Fig. 6. The histogram of each original image illustrates how the pixels are distributed by graphing the number of pixels at every color of RGB [14]. It is clear that the histogram of the scrambled image is different from the respective histograms of the original image.

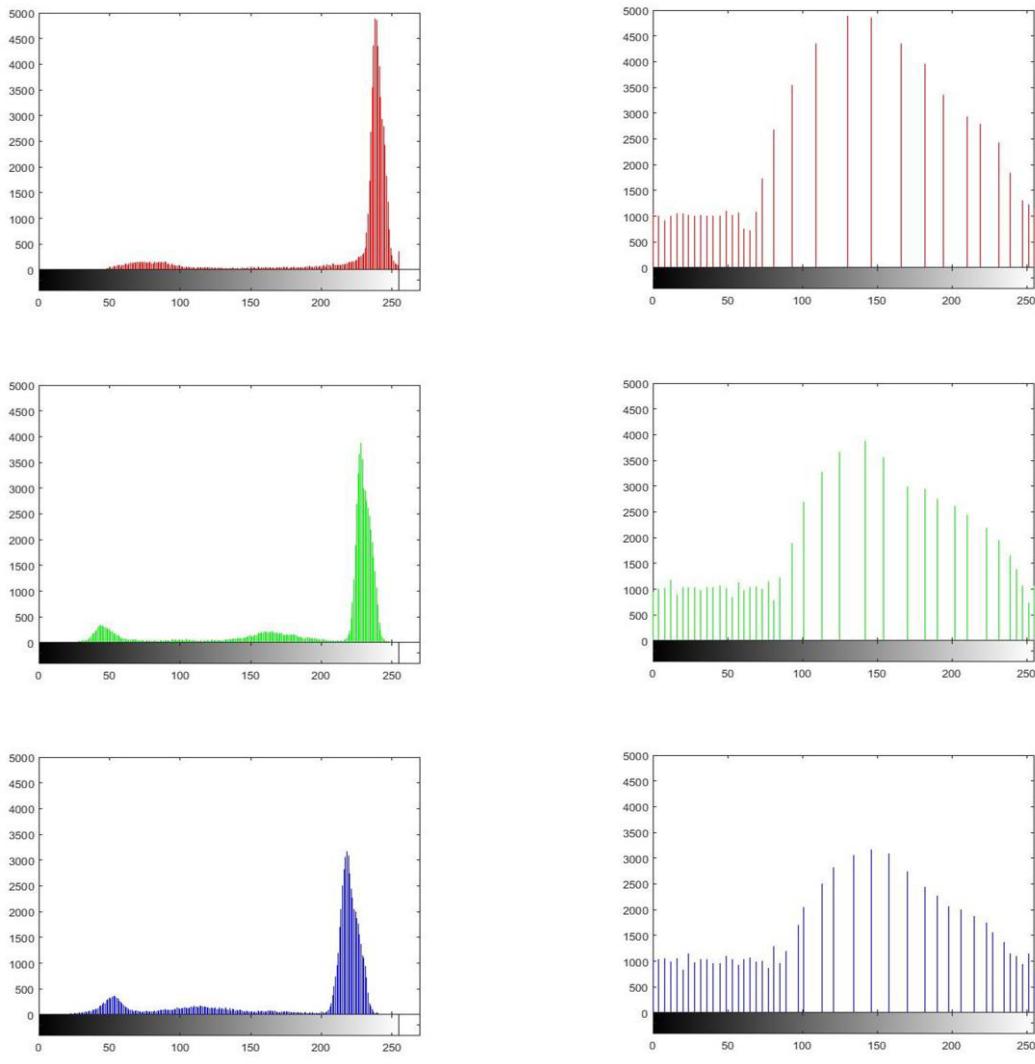


Figure 6: "Alice" image histogram in three channels RGB

(a): Original (b): Scrambled

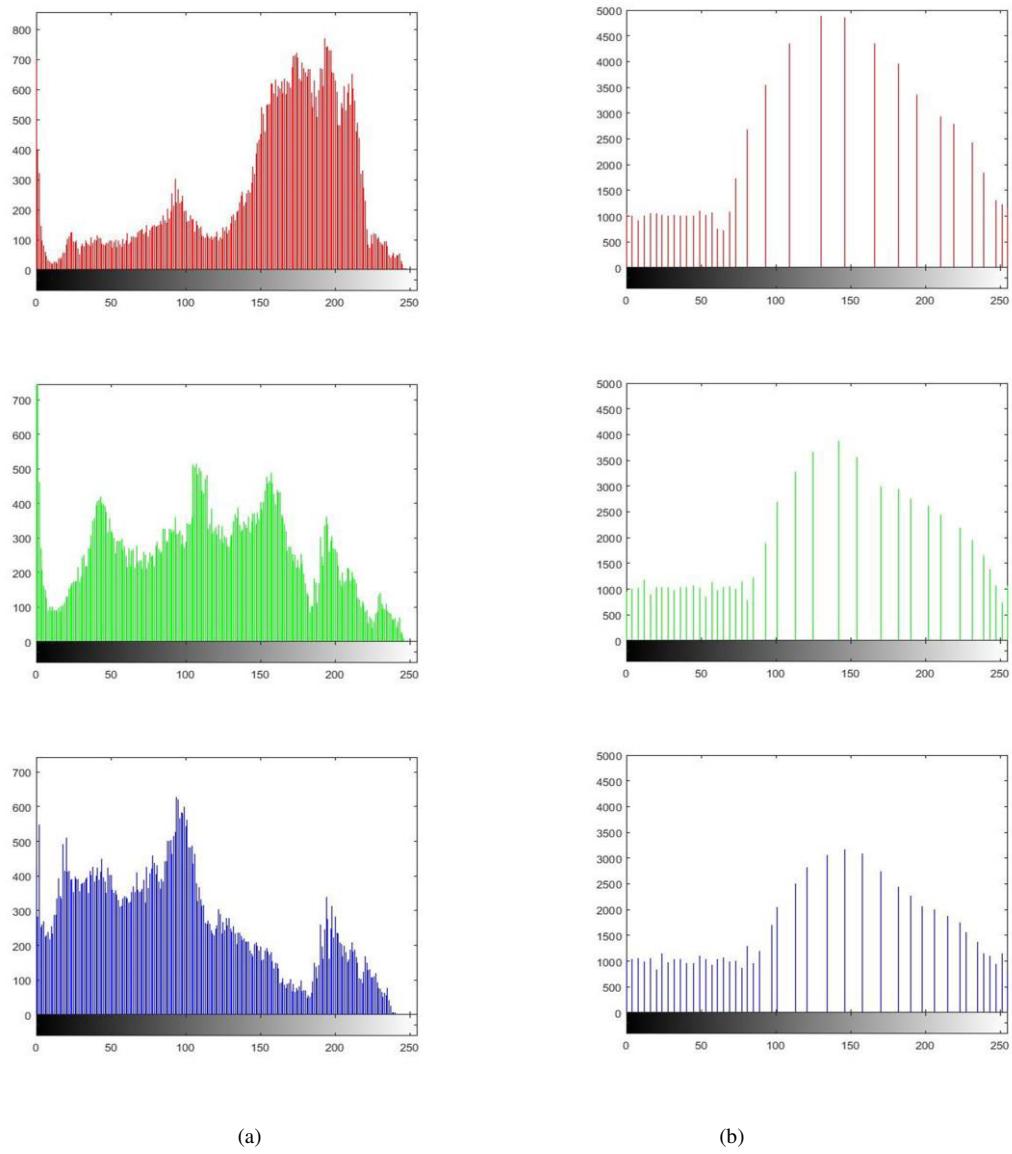


Figure 7: "Gallery" image histogram in three channels RGB

(a): Original (b): Scrambled

#### 4.2. Correlation of two adjacent pixels

In addition to the histogram analysis, we have also analyzed the correlation between two vertically adjacent pixels, two horizontally adjacent pixels and two diagonally adjacent pixels in plain image and cipher image respectively.

A correlation is a statistical measure of security that expresses a degree of relationship between two adjacent pixels in an image or a degree of association between two adjacent pixels in an image. The aim of correlation measures is to keep the amount of redundant information available in the scrambled image as low as possible [11, 15].

Equation (3) is used to study the correlation between two adjacent pixels in the horizontal, vertical, diagonal and anti-diagonal orientations:

$$C_r = \frac{N \sum_{j=1}^N (x_j \times y_j) - \sum_{j=1}^N x_j \times \sum_{j=1}^N y_j}{\sqrt{(N \sum_{j=1}^N x_j^2 - (\sum_{j=1}^N x_j)^2) \times (N \sum_{j=1}^N y_j^2 - (\sum_{j=1}^N y_j)^2)}} \quad (3)$$

where  $x$  and  $y$  are the intensity values of two adjacent pixels in the image and  $N$  is the number of adjacent pixels selected from the image to calculate the correlation. Results for the correlation coefficients of two adjacent pixels are shown in tables 1and 2.

In the experiments results, 3000 pairs of two adjacent pixels are randomly selected. Fig. 8 shows the distribution of two adjacent pixels in the original image and the encrypted-image. There is very good correlation between adjacent pixels in the image data [16, 17], while there is only a small correlation between adjacent pixels in the scrambled image.

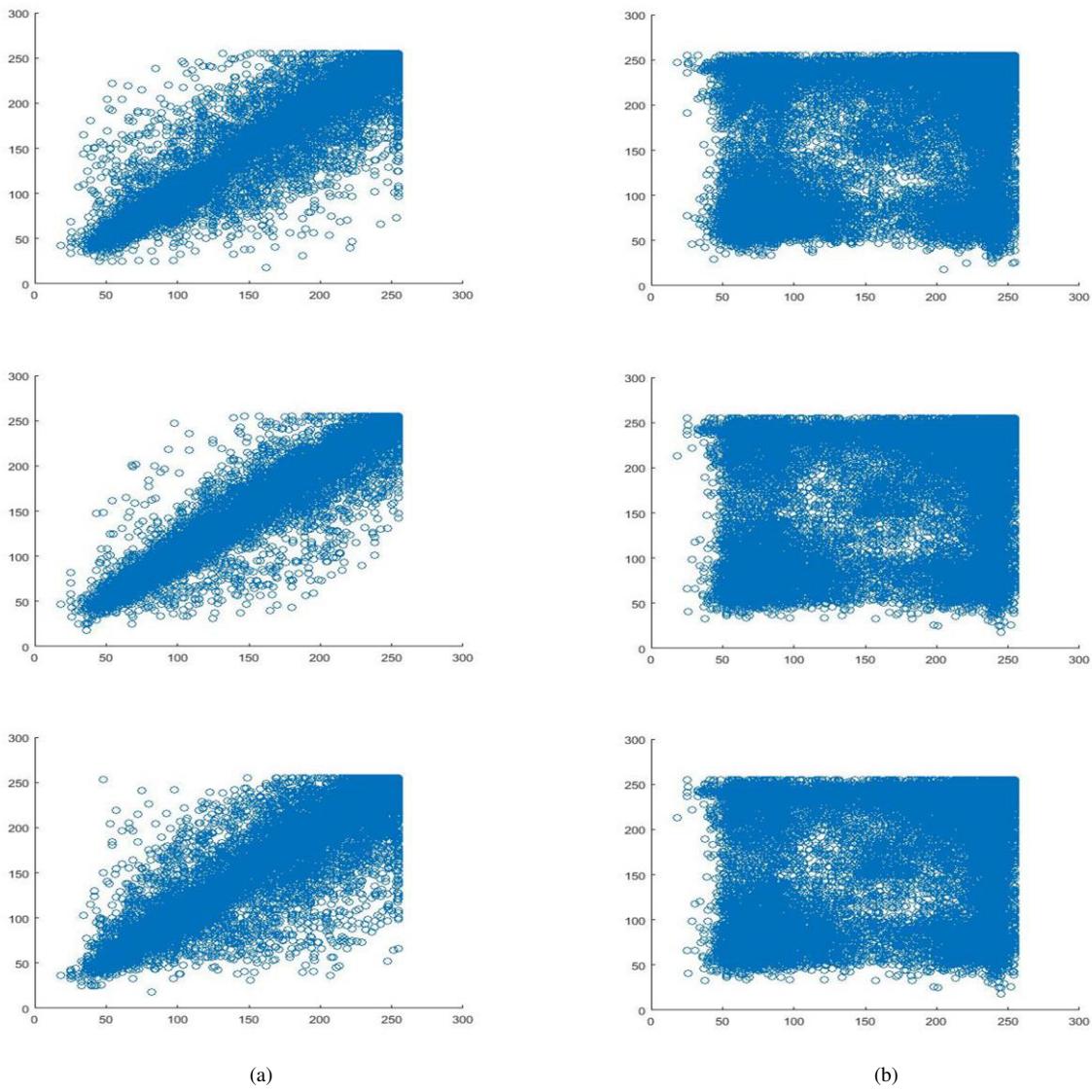


Figure 8: Horizontal, vertical and diagonal correlation of original and scrambled "Alice" image

(a): Original image    (b): Scrambled image

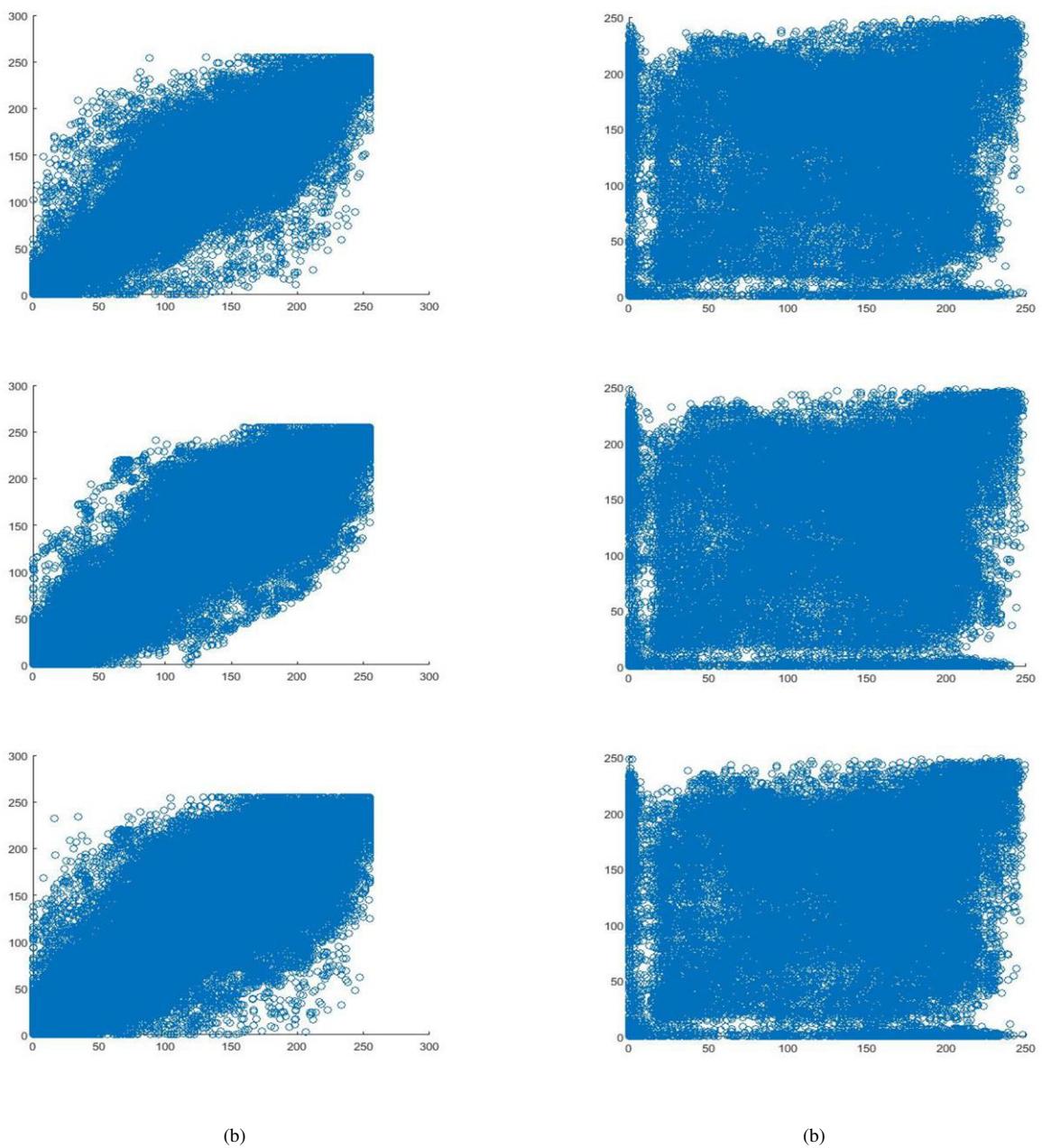


Figure 9: Horizontal, vertical and diagonal correlation of original and scrambled "Gallery" image  
(a): Original image (b): Scrambled image

Table I  
Correlation coefficient corresponding to original and scrambled images

Direction		Horizontal	Vertical	Diagonal
Alice image	Original	0.9759	0.9855	0.9677
	Scrambled	0.5632	0.5656	0.5612
Gallery image	Original	0.9878	0.9777	0.9679
	Scrambled	0.3977	0.3921	0.3832

Table II  
Correlation coefficient corresponding to original and recovered images

Direction		Horizontal	Vertical	Diagonal
Alice image	Original	0.9759	0.9855	0.9677
	Renew	0.9759	0.9855	0.9677
Gallery image	Original	0.9878	0.9777	0.9679
	Renew	0.9674	0.9665	0.9290

#### 4.3. PSNR

Peak Signal to Noise Ratio (*PSNR*) criterion is used to test the unobservable factor. This measure indicates the degree of similarity between the watermark images and a watermark images. *PSNR* is expressed mathematically in the following form:

$$PSNR[dB] = 10 \log_{10} \left( \frac{255^2}{EQM(I_o, I_R)} \right) \quad (4)$$

where *EQM* is the mean square error between the two images ( $I_o$  original,  $I_R$  recovered).

$$EQM(I_o, I_R) = \frac{1}{mn} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} (I_o(x, y) - I_R(x, y))^2$$

To recover the two images, we apply the inverse of the proposed algorithm in figure 2. The result is shown in figure 10.



Figure 10: "Alice" and "Gallery" recovered image

PSNR high means: Mean square error between the original image and reconstructed image is very low. It implies that the image been properly restored. In the other way, the restored image quality is better; in our case, the values of *PSNR* are as follow:

$$\text{PSNR (Alice)} = \text{Inf}$$

$$\text{PSNR (Gallery)} = 51.75$$

The result is much closed with the correlation coefficient.

- For "Alice", the correlation coefficient for the original and renew image are identical. The *PSNR* equal to INF, that means the renew image is identical to original image.
- For "Gallery", the correlation coefficient for the renew image is at 40% of the original image, that justify the corresponding *PSNR* value.

## V. CONCLUSION

In this paper, a new image scrambling algorithm, by using image scrambling to encrypt the image to improve the security of image. The new algorithm based on chaotic system and decomposition and recombination of pixel values is able to scramble pixel positions and pixel values of images. Analysis of the statistical information of scrambled images in the experimental tests shows that the present algorithm provides reasonable security. Owing to the strong irregularity of the sorting transformation that improves the effect of the scrambling. The experimental results show that the algorithm is effective to scramble the image and can provide high security. It simulates scrambling under Matlab 7 to confirm it.

## REFERENCES

- [1] Pawan N. Khade and Prof. Manish Narnaware, "3D Chaotic Functions for Image Encryption," IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 3, No 1, May 2012, pp: 323- 328.
- [2] P. Ping, Z. J. Wang, and F. Xu, "A two-dimensional cellular automata based method for multiple image", International Conference on Computer Science & Service System, vol. 112, pp. 101-104, 2014.
- [3] F. Chen, K. W. Wong, X. Liao, and T. Xiang, "Period distribution of generalized discrete Arnold cat map", Theoretical Computer Science, vol. 552, no. 4, pp. 13-25, 2014.
- [4] L. Chen, D. Zhao, and F. Ge, "Image encryption based on singular value decomposition and Arnold transform in fractional domain", Optics Communications, vol. 219, no. 6, pp. 98-103, 2013.
- [5] Zhu Liehuang, Li Wenzhuo, Liao Lejian and Li Hong. "A Novel Image Scrambling Algorithm for Digital Watermarking Based on Chaotic Sequences" [J]. International Journal of Computer Science and Network Security, 6(8B): 125-130, 2006.
- [6] Matthews R. "On the derivation of a 'chaotic' encryption algorithm [J]". Cryptologia, 1989, 13:29-42.
- [7] Schäfer J. Fast encryption of image data using chaotic Kolmogorov flow [J]. J. Electronic Imaging, 1998, 7(2): 318-325.
- [8] Yu X Y, Zhang J, Ren H E, Xu G S and Luo X Y. "Chaotic Image Scrambling Algorithm Based ob S-DES [J]". Journal of Physics: Conference Series, 2006, 48: 349-353.

- [9] S.N. Lagmiri, M. Amghar and N. Sbiti, "Synchronization between a new chaotic system and Rössler system by a high gain observer", IEEEXplore, May ,2015.
- [10] J. Khan, J. Ahmad, and S. O. Hwang, "An efficient image encryption scheme based on: Henon map, skew tent map and s-box," International Conference on Modeling, Simulation and Applied Optimization, vol. 10, pp. 1- 6, 2015.
- [11] Burger, W. and M. Burge, *Digital image processing: "an algorithmic introduction using Java"*. 2008: Springer-Verlag New York Inc.
- [12] D. L, "Color image encryption algorithm based on chua's circuit and chen's hyper-chaotic system," Journal of Information & Computational Science, vol. 12, pp. 1021-1028, 2015.
- [13] S. Liu, J. Sun, and Z. Xu, "An improved image encryption algorithm based on chaotic system", Journal of Computers, vol. 4, no. 11, pp. 1091- 1100, 2009.
- [14] Abderrahim, N.W., F.Z. Benmansour, and O. Seddiki, "*Integration of chaotic sequences uniformly distributed in a new image encryption algorithm.*" 2012.
- [15] Jolfaei, A. and A. Mirghadri, "*Survey: image encryption using Salsa20*". International Journal of Computer Science Issues, 2010. 7(5): pp. 213-220.
- [16] 17. H. El-din. H. Ahmed, H.M.K., O. S. Farag Allah, "*Encryption quality analysis of the RC5 block cipher algorithm for digital images*". Optical Engineering, 2006. 45( 10).
- [17] 18. Ibrahim S I Abuhaiba and M.A.S. Hassan, *Image Encryption Using Differential Evolution Approach in Frequency Domain*. Signal & Image Processing : An International Journal(SIPIJ), 2011. 2, No.1: p. 51-69.

# Data Partitioning In Cloud Storage Using DESD Crypto Technique

B.Muthulakshmi<sup>1\*</sup>, M.Venkatesulu<sup>2</sup>

<sup>1</sup>\*Department of Computer Applications, Kalasalingam University, Krishnankoil, Tamil Nadu, India.

<sup>2</sup>Department of Information Technology, Kalasalingam University, Krishnankoil, Tamil Nadu, India.

\*Corresponding Author: selvamayil2010@gmail.com

## Abstract

With the growth of cloud technologies, computing resources and cloud storage have become the most demanding online services. There are several companies desiring to outsource their data storage and resources as well. While storing private and sensitive data on a third party data center, it is necessary to consider security and privacy which become major issues. In this paper, a novel Double Encryption with Single Decryption (DESD) crypto technique is proposed to secure the data in cloud storage. The proposed technique comprises of encryption and decryption phases where in the encryption phase the data is randomly partitioned into multiple fragments. Double encryption is done on each fragment by prime numbers, as well as Invertible Non-linear Function (INF). These multiple encrypted data are stored at the multiple cloud storages with the help of cloud service provider (CSP). After all verification process the data user collects the key from the data owner and decrypts the gathered data from the cloud with the knowledge of inverse INF. The proposed crypto technique provides more security and privacy to cloud data and any illegitimate users cannot retrieve the original data. The performance of the proposed DESD technique is compared with AES and Triple DES techniques and the experimental results are plotted which shows the proposed technique is efficient and faster.

**Key words:** Cloud computing, cloud service provider, DESD crypto technique, Invertible Non-linear Function, AES and Triple DES.

## 1. Introduction

In recent years, this fast growing innovative technology offers users with several paperless services which are available online, for example, e-banking, e-billing, e-mail, e-shopping and e-transaction etc. These paperless services need data

exchange through online. This data might be any personal or sensitive information such as credit or debit card details, business secrete information, banking transactions and so on. These kinds of information need more security as disclosure of such personal data to any illegitimate user can produce extremely hazardous consequences. There is a high necessity for user's security while exchanging their personal information through un trusted networks. Thus, it is necessary to develop a security mechanism for converting user's personal or sensitive information to some other unreadable format. While sending such information it is essential to build it harder for intruders to collect some observed information. Cryptography is one of the techniques to achieve it.

In cloud computing ,user's data (i.e. data owner) is stored at some untrusted third party that needs extreme protection as data owner does not possess any physical access on the information. Data privacy and security of user or owner are consistently a vital issue in cloud computing (Dai Yuefa et al (2009), Mohit Marwaha and Rajeev Bedi (2013)). There are several advantages such as low cost and easy access on data provided by the cloud but privacy and security problems is of concern while storing user's personal and sensitive data to cloud storage (M. Mohamed et al (2013)). Data in cloud storage might be attacked in two manners such as inside or outside attack (L. Arockiam and S. Monikandan (2013)). If an attacker attempts to access the cloud data while in transition or at rest which is not legitimized, then it is known as outside attack. An attack from the cloud administrator side is defined as inside attack. When compared to the outside attack, the inside attack is

really hard to identify and the data owner or user must be very careful while storing and retrieving their personal data to or from the cloud storage. Moreover the retrieved data by the authorized user from the cloud should not be in actual format as there is high possibility of outside attack. Therefore all the data must be converted into unreadable format by encrypting before storing it in cloud; then its actual format is revert back by decryption. This should be possible with the aid of cryptography.

Cryptography is classified into two techniques namely “code making” and “code breaking”. The code making involves to covert a message or data into other incomprehensible/unreadable format to secure it from any malicious activity of malicious users whereas the code breaking provides the solution known as cryptanalysis (Chris Christensen (2006)).

The major objective of cryptography is preventing intruders from obtaining the actual data and permits only legitimate users to obtain the correct information without any modification. Utilization of cryptographic strategies guarantees the user's personal information remains secure from any changes and illegitimate users. These illegitimate users cannot break encrypted code of original information while legitimate users only have the authority to revert back the translated information into actual format (Sinkov A (1996)). The entire process of conversion of original data and reversing back the exact data is called as encryption and decryption respectively.

This paper proposed a novel Double Encryption with Single Decryption (DESD) crypto technique to protect cloud data. Data of large volume is split into number of small fragments by data partitioning process. Then each partition is subjected to encryption and here double encryption is done. The first encryption is accomplished with prime numbers. For that the prime numbers are generated randomly and the number of generated prime numbers equal to twice the number of partitioned data when the data owner wants to produce four encrypted forms. Based on the interest of data owner he/she can produce 8 or 16 or 32 encrypted forms for a single data part. Then complements of all primes are computed. So each partition is encrypted with a prime and its

complement. After the first encryption a large integer is generated and is divided into number of small integers which is equal to the number of data partitions. Each small integer is added with each encrypted data. Each resultant cipher is then subjected to another encryption using invertible non-linear function (INF) which has two random integers. The second encryption is achieved by multiplying each data partition with the first integer and added with the second integer. At user end a single decryption is enough to decrypt the data and the key is subtracted to retrieve original data where the key is a large integer value. An important thing is that the data user must have the knowledge of the inverse invertible non-linear function for decryption.

The rest of the paper is organized as follows: Section 2 presents the related works on cryptographic techniques and section 3 presents the problem definition. In section 4, the proposed method is presented in detail. Section 5 deals with the experimental results and in section 6 the paper is concluded with scope for future work.

## 2. Related work

V.Masthanamma, G.Lakshmi Preya (2015) examine about the usage of cryptography schemes, to enhance the security of encrypted data that is sent by the cloud users to cloud server. The fundamental goal is to perform encryption and decryption of data in a secured way with consumption of very less time and low cost for both encoding and decoding process. Various amounts of keys are produced and repeated attacks are observed. Thus by repeating the strategy it assists the data to remain safe against the attacks to extend the security of decoded data that is sent by the cloud users to cloud server.

H.Y. Lin and W.G. Tzeng (2012) presented a threshold proxy re-encryption scheme in which data security is accomplished using decentralized erasure code. This makes the system stronger and privacy issues of cloud service provider (CSP) are solved. Here the data is stored in a cloud storage server in encrypted format and when a user requests the data, the data holder sends the re-encryption key to the server that again encrypt the same data for requested user. The authors consider that the cloud storage comprises of storage and key servers where in

storage server the data storing operation are performed. In order to decrypt the encoded and encrypted data with n codeword symbols, each key server has to independently perform partial decryption alone.

Pancholi et al (2016) have presented the method of using diverse parts by the ciphers and its converse for all purposes eliminates the keys in AES that is the drawback of DES. In AES, the likelihood of proportionate keys is taken out for nonlinearity of the key extension for all purposes. For several microcontrollers an implementation correlation among AES, DES and Triple DES exhibits that AES and Triple DES require a PC expense of the same request. Another execution evaluation reveals that AES has great status over estimations 3DES, DES and RC2 to the extent of execution time with different package size and throughput for encoding and furthermore decoding. Similarly changing data order, for instance, image instead of content, it has been discovered that AES possess a benefit over Blowfish, RC2 and RC6 with respect to time usage.

K. Nasrin, et. al. (2014) dealt with cloud storage framework which is the most important research area in cloud computing in which security is considered as one of the vital concerns. The authors combined the asymmetric and symmetric key approaches utilizing AES and RSA algorithms and derived a novel mechanism. AES is useful for key sharing and less overhead cryptographic technique and complex phenomena is created by RSA to provide security from attackers. The main attention of the attackers was on demonstrating secure data communication from defenseless or vulnerable networks.

Jayant, D. et al. (2015) presented a novel mechanism called role base access control by applying AES and RSA algorithm for providing a secure communication environment for open cloud environments. The authors used RSA and AES algorithms for the purpose of encryption and decryption where access control is achieved using RBAC mechanism. According to the model of RBAC the uploading rights and several rights to several users were given.

In this paper, a novel DESD crypto technique is proposed to provide privacy and security for

confidential and sensitive data stored in cloud server. It requires less computation time with low cost. It also provides better protection against intruders and malicious activities with faster operations.

### 3. Objective and issues

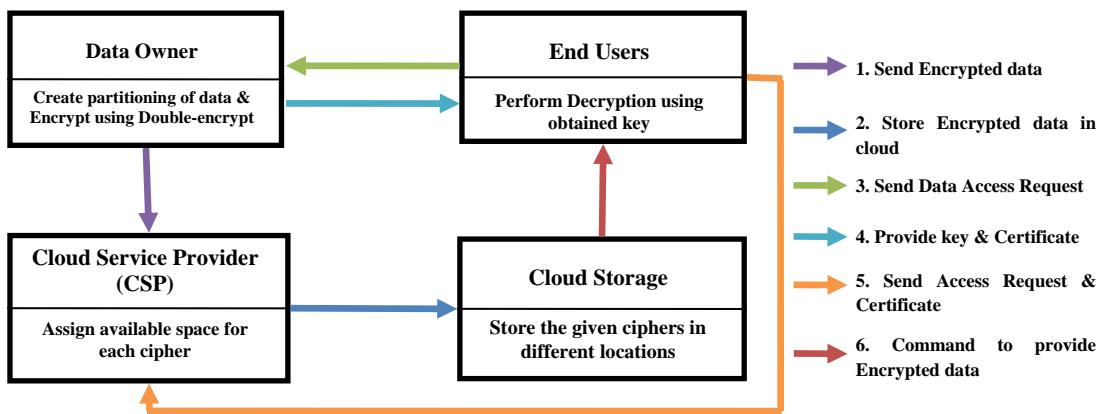
The main objective of this paper is to design an efficient cryptographic technique which is simple and consumes less time to perform encryption and decryption operations on data stored in cloud. The encrypted data should require limited space for storage. Some of the following privacy and security issues are rectified.

- **Access control:** Failure of CSP may happen at some situation on cloud environment that leads the chances of intruders and malicious activities.
- **Lack of user control:** In cloud user data is stored at some remote location and its complete control is taken by CSP i.e. the user has no control on its data.
- **Control policy:** The CSP may have self-interest on user's data at some network conditions. Thus it is necessary to implement security mechanism for CSP to provide control policy in the cloud environment.

## 4. Proposed Methodology

### 4.1. System Model

The proposed system model comprises of four entities such as data owner, cloud service provider (CSP), cloud storage and data user. Figure 1 illustrates the system model of the proposed work and its flow of operation is also explained.

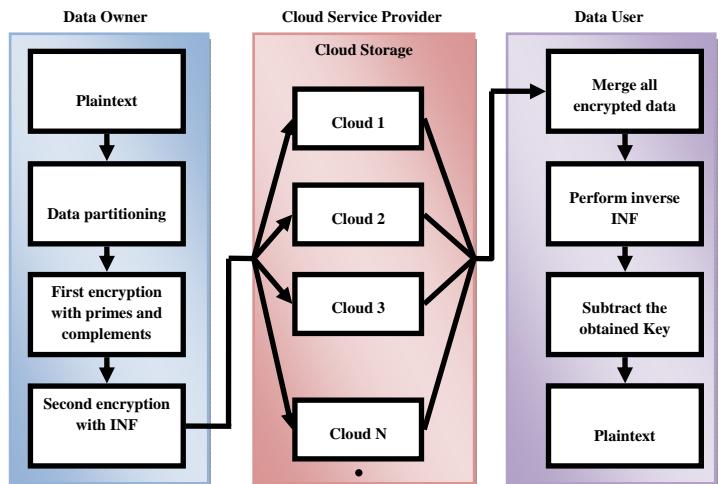


**Figure 1: Proposed system model**

The operation of **data owner** is to collect data in any form like image, text, audio and video which will then be building a document index which is partitioned into several numbers of small fragments. Each fragment is encrypted multiple times and outsourced to the cloud storage. The **cloud service provider (CSP)** is responsible to allocate available space for outsourced data at different storage location of a single cloud or different clouds. The CSP has the complete control on the cloud storage i.e. once the data is stored in the cloud its complete control is taken by the CSP. Here 'N' number of clouds is taken to store the user's data. In **cloud storage**, the ciphers are stored in allocated storage space. If the **data user** wants to access data in cloud he/she must be verified by the CSP to verify his/her authorization. If he/she is an authorized user then it allows sending data access request to the data owner. The data owner responds the request by sending authentication certificate with a decryption key. By verifying the gathered authentication certificate the CSP command the storage to provide data. Finally the decryption is done by the collected key from the data owner.

#### 4.2. Detailed contribution

The detailed contribution of the proposed work is explained through the block diagram demonstrated in Figure 2. This block diagram comprises of three major blocks such as data owner, CSP and data user. Each of its operations is explained below in a detailed manner.



**Figure 2: Block diagram**

The data owner comprises the data in the form of plaintext which is large size. So it is partitioned into multiple small fragments. There are several advantages in data partitioning such as: 1) Processing of large volume of data makes the operation complex. 2) Uploading and downloading of these small fragments requires relatively very less time. 3) These are very easy to access. Then each partitioned fragment is encrypted with prime numbers and its complements for multiple times. With the help of Inverse Non-linear Function (INF) second encryption is accomplished which produces ciphers with unreadable format. These are outsourced to the different locations at same cloud storage and different

cloud storage. For that the CSP has to process these data to allocate available space for storage. After all verification process the CSP command the cloud to provide the data stored at different locations. The data user merges all the collected encrypted data and performs inverse INF. Finally to get the plaintext the key is subtracted from the previous results.

Symbol	Description
$X_i$	Plaintext
$i$	Number of split or partitioned data $i = 1, 2, \dots, n$
$P_N$	Number of generated primes
$P_i, Q_i$	Set of prime numbers
$P_i^c, Q_i^c$	Prime complements
$p$	Bit of the prime. Here $p = 32$
$D_i$	Random Integer
$d_n$	Number of split of random integer
$f_i$	Generated ciphers
$S_i$	Sum of encrypted form

### 4.3. Encryption

The large volume of data is to be stored in the cloud effectively. So the large volume of data is partitioned into number of small partitions or fragments at the first step of encryption. Here double encryption algorithm is proposed to encrypt each partition. The first encryption is done with the prime numbers and its complements. The pseudo plaintexts or cipher texts are obtained by the second encryption with Invertible Non-linear Function (INF) and its general form is given as,

$$g(x) = ax + b$$

where  $a$  and  $b$  are integers and  $y$  denotes cipher texts obtained through the first encryption. Each cipher part is multiplied with  $a$  and then added with  $b$ . These ciphers are stored at different locations of a single cloud or multiple clouds.

#### a. Double Encryption Algorithm

Input :  $X_i$

Method:

- i. Random partitioning  

$$X_i = \{x_1, x_2, x_3 \dots x_n\}$$
- ii. Generate  $P_N$  and  $P_N = 2X_i$
- iii. Take  $P_i, Q_i$
- iv. Compute  $P_i^c, Q_i^c$   

$$[\because K_i^c = 2^{p+1} - K_i]$$
- v. Generate  $D_i$  and split it into small integers  

$$D_i = d_1, d_2, \dots, d_n$$
- vi.  $d_n = x_n$  (Here, we take  $n = 4$ )
- vii. First encrypted data  

$$y_1 = (x_1 * P * Q) + d_1$$

$$y_2 = (x_1 * P * Q^c) + d_2$$

$$y_3 = (x_1 * P^c * Q) + d_3$$

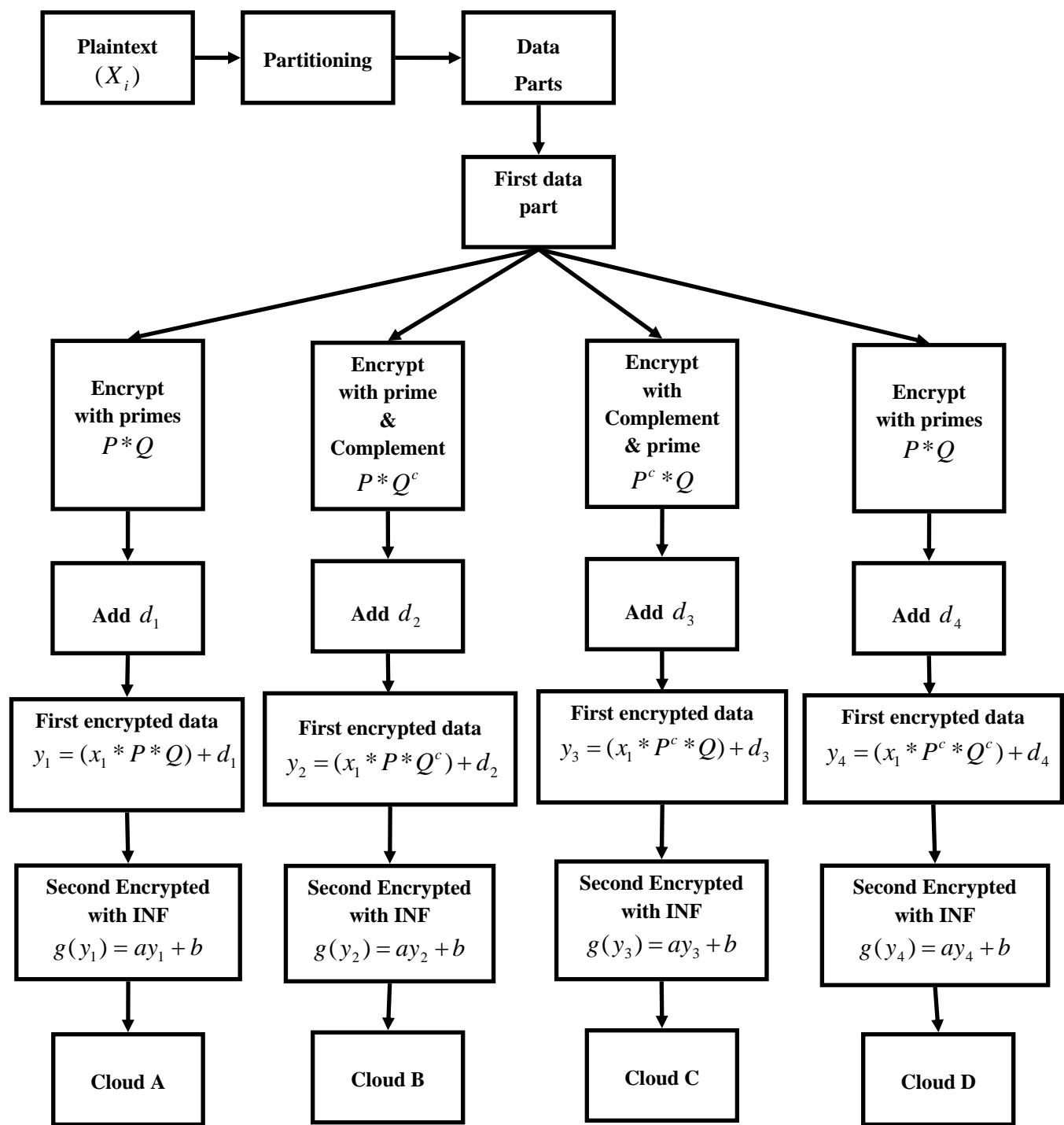
$$y_4 = (x_1 * P^c * Q^c) + d_4$$
- viii. Second encrypted with INF  

$$g(y_1) = ay_1 + b$$

$$g(y_2) = ay_2 + b$$

$$g(y_3) = ay_3 + b$$

$$g(y_4) = ay_4 + b$$
- ix. Apply the above steps on each data part (up to  $x_n$ ) and store the obtained multiple ciphers in different locations of a single cloud storage or different cloud storages.



**Figure 3: Flow diagram of encryption algorithm**

#### 4.4. Decryption

In most of the cryptographic techniques, decryption keys are included in the encrypted data stored in cloud storage. But in our proposed crypto technique the data owner has the decryption keys which are given to the requesting authorized data users by the data owner with an authentication certificate. After verifying this certificate the CSP command the storage to provide cipher text or encrypted data. With the knowledge of inverse INF and using decryption key the user decrypt the encrypted data. The general form of inverse INF is as follows,

$$g^{-1}(x) = \frac{x-b}{a}$$

where a and b are integers and x is the cipher text.

##### a. Decryption Algorithm

Step 1: Apply inverse form of INF

$$G_i = g^{-1}(y_i) = \frac{y_i - b}{a}$$

Step 2: Add all first encrypted ciphers

$$S_i = f_1 + f_2 + f_3 + f_4$$

Step 3: Subtract large integer

$$Z_i = S_i - D_i \quad (Z_i \text{ with padded zeros})$$

Step 3: Delete the padded zeros

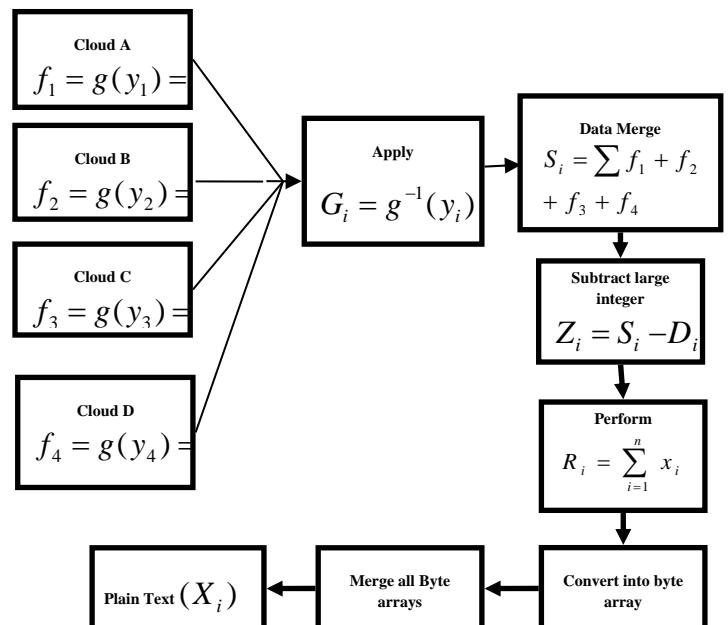
Step 4: Perform the above steps on all the four encrypted data parts(up to  $x_n$ ) and sum all of them.

$$R_i = \sum_{i=1}^n x_i$$

Step 4: Convert into byte array

Step 5: Merge all Byte arrays

Step 6: Get original plaintext ( $X_i$ )



**Figure 4: Flow diagram of decryption algorithm**

#### 5. Performance Analysis

In this section, the performance of the proposed DESD crypto technique is analyzed and compared with existing techniques in a detailed manner. As we know that the privacy and security are the most important concerns in cloud computing. All existing cryptographic techniques tried to provide privacy and security to the cloud storage at its level best. There are thousands of cryptographic techniques proposed previously and we cannot take all of them for comparison. So we take two standard

cryptographic techniques among them for comparison such as AES and Triple DES.

Consider a situation that an intruder gathers data from cloud by breaking protection mechanism of CSP. We compare the performance of existing and proposed techniques at this situation.

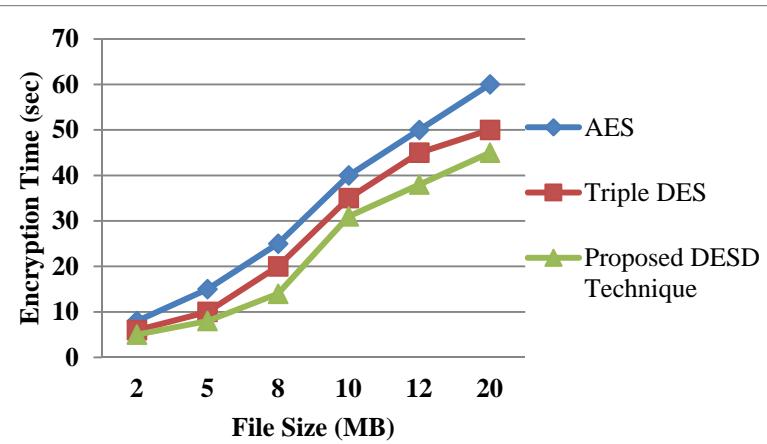
First we take existing AES and Triple DES techniques which provide complete control on data to CSP. So the intruder can easily break the gathered encrypted files since they are encrypted with decryption keys. Moreover the existing techniques cannot give cent percentage privacy assurance where the data encryption is done by the CSP. If the CSP is self-interested on its data it can misuse the data without knowing the data owner.

In our proposed technique, the data owner has the complete control on data by keeping the decryption key with him/her and they store encrypted file only at the cloud storage. Without the knowledge of inverse INF and decryption key the intruder cannot decrypt the file and retrieve the data. Therefore it provides complete access control on user data. The self-interest of CSP on data comes under the control policy which is the most significant issue in cloud environment. Here the only task of CSP is to allocate storage space for data and it never involves in data partitioning and encryption. Thus this self-interest cannot affect the cloud data. From this we can summarize that the proposed DESD crypto technique is much secure and provide better privacy to cloud users.

## 6. Result and Discussion

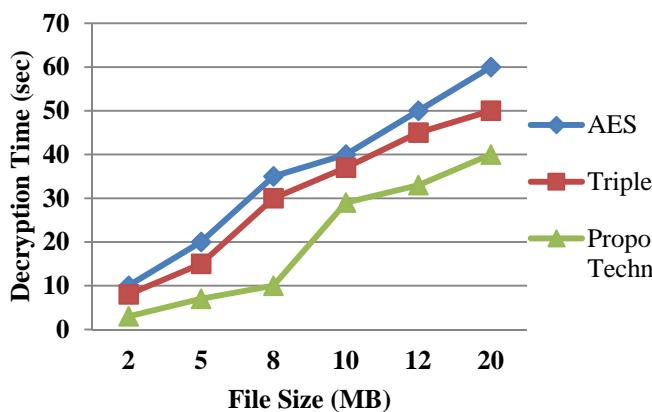
The experiment is conducted using Intel(R) Core(TM)2 Duo CPU processor with 4 GB RAM and on Windows 7 platform. The experiment was implemented using Java programming. In order to prove the efficiency of the proposed crypto technique it is compared with some other existing cryptographic

techniques. The proposed DESD technique is compared with AES and Triple DES techniques. In our implementation we employed same size of input files and examined the performance of all three techniques. Here the encryption time and decryption time is compared against file size.



**Figure 5: File size (MB) vs. Encryption Time (sec)**

In Figure 5, the encryption time of each file size is plotted for AES, Triple DES and proposed DESD technique. Generally the complex operations required more time to process the data. But the operations of both encryption techniques in our proposed crypto technique are simple and easy to process the data. So the time to encrypt different file size is reduced when compared to other techniques. From the graph it is clearly shows that the proposed DESD technique possesses less encryption time than AES and Triple DES.



**Figure 6: File size (MB) vs. Decryption Time (sec)**

Figure 6 demonstrates the comparison plot of file size (MB) and decryption time (sec) for proposed and existing techniques. This uses the same file size that of encryption. The time required to decrypt the encrypted data is known as decryption time. As we mentioned above the proposed technique is simple i.e. it required simple mathematical operations to encrypt data files. From that we can know that the decryption is also a simple process. Moreover a single decryption is enough to decrypt the encrypted data which is encrypted for a couple of time. Thus the decryption requires very less time. From figure 6, it can be observed that the proposed decryption requires very less time compared to AES and Triple DES techniques.

From the above two comparisons it is proved that our proposed DESD technique is efficient and faster by its simple operation and it is much secure because it never enclose the decryption key with the encrypted data. Also the data user has to possess knowledge on inverse INF and he/she must communicate to the data owner to get the decryption key. Hence the data is protected against intruders, unauthorized users and self-interest of CSP.

## 7. Conclusion

Data privacy and security are considered as the most important issues in cloud data storage. Though cloud can provide easy and flexible data storage, but there are possibilities for intruders and malicious activities. In cloud, the stored data may be confidential which requires more security concerns. In this paper, we

proposed a novel Double Encryption with Single Decryption (DESD) crypto technique for secure data storage in cloud. Data partitioning is done to make the storage easy and effective which also provides flexible data access with less storage cost. Then double encryption is performed on each partitioned data which includes two encryptions namely encryption with prime numbers, as well as its complements and then with an INF encryption. Using the proposed decryption algorithm the obtained data can be decrypted by the user. The major benefit of this proposed technique is, the encryption is done by the data owner and the encrypted data is only stored at the cloud storage with the help of CSP. The authorized users have knowledge on inverse INF which is another important factor for decryption. Thus the intruders and third parties aren't able to retrieve and misuse the cloud data without knowledge on inverse INF and decryption key. In experimental section the proposed technique is compared with AES and Triple DES techniques. The performance analysis is done using some parameters such encryption time and decryption time against file size. From the Figure 5 & 6, it is clearly observed that our proposed crypto technique is efficient and faster in terms of reduced encryption and decryption time compared to other techniques. In future the proposed DESD crypto technique will be used to encrypt video files.

## References

- [1] Dai Yuefa, Wu Bo, et al. "Data security model for cloud computing." Proceedings of the 2009 International Workshop on Information Security and Application (IWISA 2009) Qingdao, China. 2009., pp 141-144.
- [2] Mohit Marwahe, Rajeev Bedi, "Applying Encryption Algorithm for Data Security and Privacy in Cloud Computing", International Journal of Computer Science Issues. Vol 10, Issue 1, January 2013, pp. 367-370
- [3] Eman M.Mohamed, Hatem S.Abdelkar and Sherif El-Etriby, "Data Security Model for Cloud Computing", the twelfth International Conference on Networks, 2013, pp. 66-74
- [4] L. Arockiam, S. Monikandan "Data Security and Privacy in Cloud Storage using Hybrid Symmetric Encryption Algorithm", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, Issue 8, August 2013, pp 3064-3070.
- [5] <http://www.nku.edu> (Fall 2006 Chris Christensen)

[6] Sinkov A., "Elementary Cryptanalysis – A Mathematical Approach", Mathematical Association of America, 1996.

[7] V. Masthanamma, G. Lakshmi Preya," An Efficient Data Security in Cloud Computing Using the RSA Encryption Process Algorithm", International Journal of Innovative Research in Science, Engineering and Technology, Vol.4,pp.1441- 1445,2015.

[8] Hsiao-Ying Lin, Member, IEEE, and Wen-Guey Tzeng, Member, IEEE"A Secure Erasure Code Based Cloud Storage System With Secure Data Forwarding" IEEE Transactions On Parallel And Distributed Systems, Vol. 23, No. 6, June 2012.

[9] Pancholi, Vishal R., and Bhadresh P. Patel. "Enhancement of Cloud Computing Security with Secure Data Storage using AES." International Journal for Innovative Research in Science and Technology 2.9 (2016):18-21.

[10] Nasrin Khanezaei, Zurina Mohd Hanapi "A Framework Based on RSA and AES Encryption Algorithms for Cloud Computing Services" IEEE Conference on Systems, Process and Control (ICSPC 2014), 12 - 14 December 2014, Kuala Lumpur, Malaysia.

[11] Bokefode Jayant D, Ubale Swapnaja A, Pingale Subhash V, Karande Kailash J., Apate Sulabha S., "Developing Secure Cloud Storage System by Applying AES and RSA Cryptography Algorithms with Role based Access Control Model" International Journal of Computer Applications (0975 – 8887) Volume 118– No.12, May 2015.

## Design a Hybrid Technique Based new Genetic approach for text Encryption

Shatha Abdullah M.Ramadhan  
Dept. of Software Engineering  
College of Computer Sc. &  
Math., University of Mosul.  
Mosul, Iraq.

Hiba Muneer Yahya  
Dept. of Software Engineering  
College of Computer Sc. &  
Math., University of Mosul.  
Mosul, Iraq.

Ruah Mouad Alyas  
College of Computer Sc. &  
Math., University of Mosul.  
Mosul, Iraq

**Abstract-** This paper presents an Applying new Genetic approach, which is Volvox reproduction(VR) algorithm using the natural concept. Then design a hybrid technique(HVR), in order to generated a wide colony of positive integer numbers, which used to encryption(decryption) of English text with different sizes. This suggested using shifting of places of letters. Where the HVR process was built using Matlab, which attend 100% success.

**Keywords-**Volvox Reproduction Algorithm(VR); HybridVR Algorithm(HVR); Genetic Algorithm; Volvox Colony; Encryption; Decryption.

### I.INTRODUCTION

In addition to the technological development that is taking place in our time, the development of mobile phones and the emergence of many of the services offered by us, the most common and the most used SMS text messaging, Which has become a means of transmitting information and personal data task (bank account data, secret army information, user information for some programs The task of the login code and password). Because this information may be stolen and hacked. The purpose of the research is to build an application that works on mobile devices supporting the Android platform, which will benefit from the SMS service provided by mobile devices to send user data securely such as access information to bank accounts and payment cards or any confidential and important information we would like to send and to convert it into a set

of symbols, letters and numbers that are not understood by suggesting a new encryption algorithm based on the phone number of the sender and receiver in the secret key between the two parties.

### II. BACKGROUND

Encryption is defined as the process of converting clear information into unintelligible information to prevent unauthorized persons from accessing or understanding information. Encryption, therefore, involves converting plain text into encrypted text. The confidentiality of information is maintained by means of methods or algorithms that have the ability to convert that information into a mixture of symbols, numbers and unintelligible characters and then transfer them via means of transport to the sender to rephrase them to their understandable form again " [5],[6],[8]"

The components of the encryption system: Plain Text, Cipher Text, Key. Encryption has become very important especially since the beginning of the twentieth century in the mid-seventies used encryption in communications and correspondence military and diplomatic and security and reached the areas and applications and other uses, including:

1. In industry and commerce.
2. In video broadcasting.
3. In banks.
4. In computer networks and personal computers.
5. In the protection of telecommunications from the capture and eavesdropping and knowledge of the secrets of others.

Strong encryption system must achieve the following objectives and characteristics: [7]

- |                                |                   |
|--------------------------------|-------------------|
| *. Privacy or Confidentiality. | *. Certification. |
| *. Data Integrity.             | *. Time Stamping. |
| *. Entity Authentication.      | *. Witness.       |
| *. Message Authentication.     | *. Receipt.       |
| *. Signature.                  | *. Confirmation.  |
| *. Authorization.              | *. Ownership.     |
| *. Validation.                 | *. Anonymity.     |
| *. Access Control.             | *. Revocation.    |

Modern key management systems Adopt two fundamental structures, the symmetrical encryption(SE), and asymmetrical encryption(AE):

Where SE is these method that encrypt the data which be sent between two entities depending on a single key. For this aim the entities firstly, will be agreement on a key, which later is used to encrypt and decrypt the data. But the asymmetrical Encryption is these method that encrypt the data which be sent between two entities depending on a multi key, as following:

first key which used to encrypt data.

second key which is used to decrypt data.

The genetic algorithm(GA) is defined as an artificial Intelligent technique that can be used to solve and modified Difficult problems.

GA considered as numerical optimization algorithms taken from the concept of natural selection and genetics.

The GA can be applied in wide and multiple fields

The genetic algorithm is successfully applied to find an acceptable (near to ideal) solution in matters Related to science, including medical and engineering sciences, as they have greatly reduced time and effort Required by system and software designers[10]

Genetic algorithm is used in the production of new generations in which have characteristics identical to the original or better ones, Sexual and non-sexual reproduction is also used in the formation of new individuals as well.

### III. LITERATURE REVIEW

There are many studies in this area, where all that studies explained the reviewing of techniques that was implemented and analyzed, the most of which is concerned with algorithms of encryption a message in the cell phone"[1],[2], [3],[4]".

### IV. THE PROPOSED METHOD

In general this work adopted the idea of the natural reproduction of Volvox algae to build a new hybrid algorithm(HVR), which is applied in security of data. Starting with the phone numbers of the sending person and the receiver in the encryption and decryption of text messages. The secret key is generated based on these numbers after a series of mathematical operations on the two numbers, so that if another phone number is different from the phone number of the receiving party, the message resulting from the decryption process will be completely different from the original message.

### V. THE NATURAL REPRESENTATION

Algae is one of the oldest and most important living organisms have been found fossils millions of years ago, and these algae still amaze humans with great benefits and new discoveries that can be performed, their forms and functions and different types and characteristics of each of them and their livelihood and importance and how to identify them and each of these aspects is fertile field For research and reflection. Because algae are of great environmental importance as they are the primary product and the first link in the food chains in water, the two methods of Reproduction, which are Eucalyptus reproduction, Sexual reproduction.

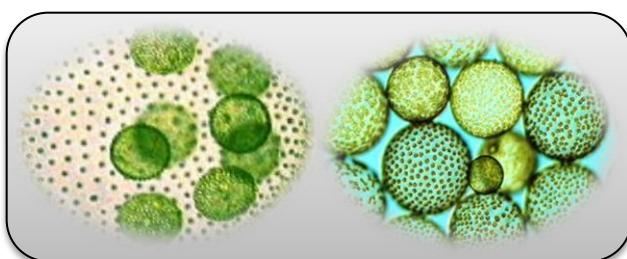
In this area, the concept was adopted Volvox's algae in this field, because The characteristic of algae is that they are one or double dwelling, as well as the characteristics listed in the following paragraphs.

Volcanic algae is one of the species of single-cell algae, the most advanced in the series of species that form spherical colonies, "see figure(1)".

Each adult colony of the Volvox consists of a huge number of whip cells, ranging from 500 to 60,000 cells [9].

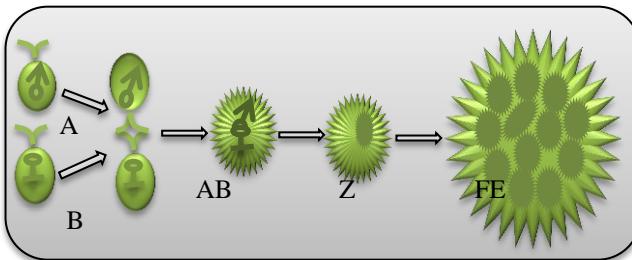
The Classifications of Volvox according to the International Classification of Plants which are:

<b>*Kingdom:</b>	<b>*Division:</b>	<b>*Class:</b>
Algae	chlorophycophyta	chlorophyceae
<b>*Order:</b>	<b>*Family:</b>	<b>*Genus:</b>
volvocales	Volvocaceae	Volvox



**Figure 1, Volvox colony**

The reproduction (Sexual reproduction), starting when the contents of the antheridia (sexual male, which are the anniversary container) are divided into a large number of male saplings. The male swimmers are released from the anhydrides and swim in the water until they reach the ( sexual female) in the egg to get the fertilized egg(Zygote). One male spore injects them to form the ulcer. The ulcers then secrete themselves as a thick membrane to resist inappropriate conditions. When conditions improve, they begin to divide the muse, followed by several simple splits to form a new colon colony, "see Figure(2)"



**Figure 2 , The Volvox Reproduction**

## VI. The Mathematical representation

The proposed method explained by that process of VR algorithm to generated a wide colony of integer numbers, which will be used as the public key to encryption (or decryption) a given text, see Figure(3) and Figure (4).

### First: The VR Algorithm

This paragraph describe a new GA, which is a Volvox Reproduction(VR) algorithm, in general has the following terminology:

\*. Where the antherizods(A) represents the male reproductive segment(First integer number).

\*. The oogonium(B) represents the female reproductive segment(Second integer number), growing in size and turning into a female quail called the egg.

\*. Combine mature (A) with (B) mature to recombination operator (crossover), which called a zygot(Z) at the first stage of nuclear fusion phase(recombination operator). Second stage turn the zygote to fertilized egg(FE), which is equivalent (The first arithmetic operations).

Finally, (Cleavage stage or multiple nuclear divisions) The fertilized egg (FE) is divided several times in succession(Frequent beating process), in order to form the new colony(NC)( Public key)

### Second The HVR Algorithm

The Hybrid Volvox Reproduction (HVR) Technique, explained by steps explain that process of VR algorithm to generated a wide colony of integer numbers, which is used as public key to encryption or decryption the given text.

The following steps are explaining that VR algorithm for generating the public key(NC) for encryption(decryption), for any given text, with any size.

Starting the algorithm with the following inputs:  $a, b, d, T, q = 3$ ,

Step(1): Find the size of the text.

**Input:** T

**Output:** L, NO. satisfying:

$$Lq_i = 2^q \cdot (2 \cdot i + 1) + 1,$$

$$Rq_i = 2^q \cdot (2 \cdot i + q)$$

$$\left\{ \begin{array}{l} \text{else if } L \leq 40 \text{ then set } NO = 2 \cdot q, \\ \text{if } L_i \in [Lq_i, Rq_i] \text{ then set } NO = 2 \cdot i + 1 \\ \forall i, i = 1, 2, \dots, N \end{array} \right\}$$

Step(2): Growth A and B as following:

**Input:** a, b, d

**Output:** SN, RN

$$A = (a)^g,$$

$$B = (b)^g, g \text{ is +ive and integer No.}$$

greater than or equal 12, and Cut d of digits from A(send number) and B(receive number),

$$i=0;$$

**Do** "The Reproduction process"

$$i++;$$

Step(3): Applying the shifting condition.

Satisfying Cheek condition of digits of A & B, as following,  $\forall i, i = 1, \dots, d$ :

$$\begin{aligned} \text{if } (A(i) == 0 \rightarrow \text{set } A(i) = 9) \text{ or } (A(i) \geq 10 \rightarrow \text{set } A(i) = A(i) - 9) \end{aligned}$$

$$\begin{aligned} \text{if } (B(i) == 0 \rightarrow \text{set } B(i) = 9) \text{ or } (B(i) \geq 10 \rightarrow \text{set } B(i) = B(i) - 9), \end{aligned}$$

**and Reverse B**, Select antheridium, say that SN=A, augonium, say that RN=B, "see table 1".

Step(4): Initial Fertilization step.

**Input:** SN, RN

**Output:** Z

$$\forall i, i = 1, \dots, d, Z(i) = SN(i) * RN(i)$$

Step(5): Second Fertilization step to get the Zygote, which is denoted by public key(PK).

**Input:** Z, NO

**Output:** NC

$$\begin{aligned} \text{Zigote}(i) &= (Z(i))^{NO} * SN(i), \forall i, \\ i &= 1, 2, \dots, N \\ NC &= (\text{Zigote}(1), \text{Zigote}(2), \dots, \text{Zigote}(N)) \end{aligned}$$

}while( $L_i \in [Lq_i, Rq_i]$ )

Step(6): Cypher operation depending on the ASCII code.

**Input:** T, NC.

**Output:** NT.

Cypher operation depending on the mathematical condition on NC as following:

$$\text{either: } NT(i) = ASCII(T(i)) - NC(i)$$

$$\text{or: } NT(i) = ASCII(T(i)) + NC(i)$$

**Note That:** Ending the algorithm with the new text(NT). If NT is plain text then this process is Encryption process, else will be Decryption process, "see Figure (4)".

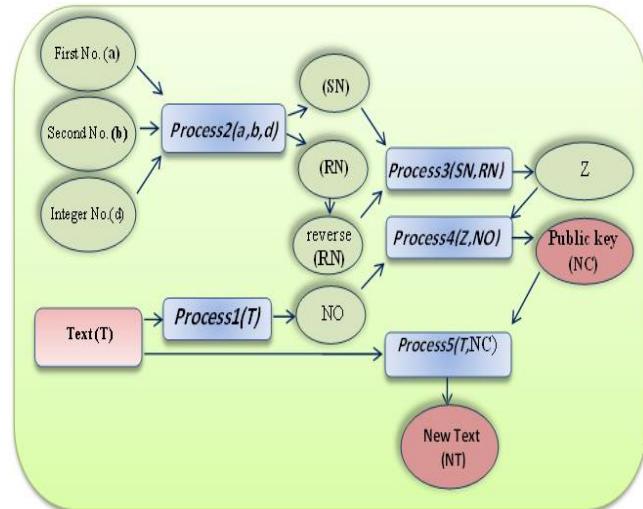


Figure 3, The VR Algorithm

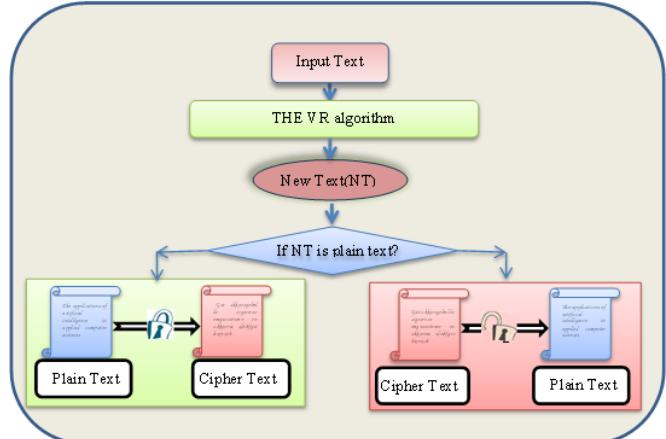


Figure 4, the HVR Method

## VII. RESULTS

The proposed method was applied to Table 1, contain Range of text ( $L_i$ ) and positive & integer number (NO), that will be using to apply the HVR algorithm for input text (T).

Table 1, Rang of  $L_i$  and NO.

$L_i:$	NO:	$L_i:$	NO:
01- 040	6	121 – 136	30
041- 056	10	137 – 152	34
057 - 072	14	153 – 168	38
073 - 088	18	169 – 184	42
089 - 104	22	...	...
105 - 120	26	...	...

Different sizes. The time of execution was measured by (mm). The program Was written by using MATLAB, to programming the scalars that measuring the Success of the HVR method. Time is the first scalar, which was Using to describe the results of Table(2).

The figures (5) and (6) Respectively are shown that, where the Symbols as following:

TE : Time of Encryption of text

TD : Time of Decryption of text

TT : Total Time of text

TTE: Total Time of execution the Encryption program

TTE: Total Time of execution the Decryption program

TT : Total Time of execution the Encryption Decryption program

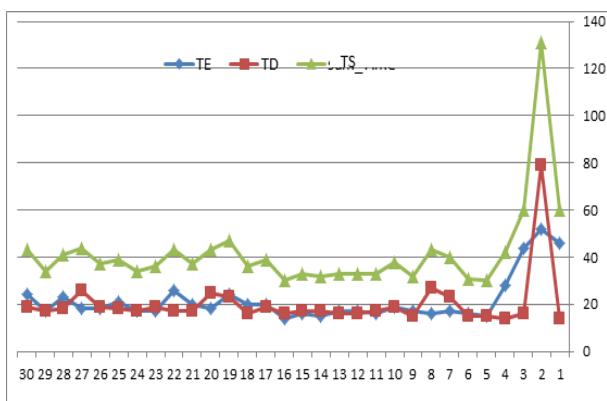


Figure 5, Time of Encryption and Decryption

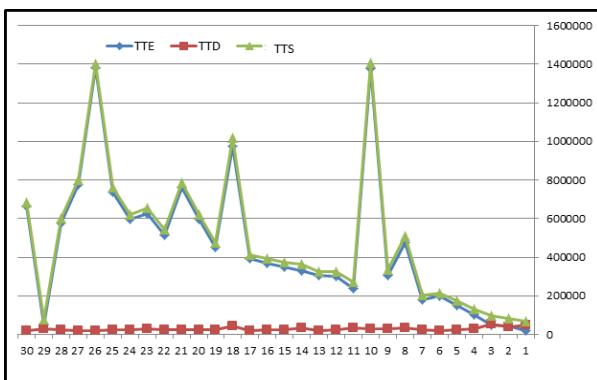


Figure 6, Total Time of Encryption and Decryption

Table 2, Time of Encryption, Decryption and Total

$L_i$	Time of partial execution			Time for Total execution		
	TE	TD	TS	TTE	TTD	TTS
1	46	14	60	21392	46610	68002
2	52	79	131	44524	38007	82531
3	44	16	60	46434	51125	97559
20	28	14	42	100165	31288	131453
40	15	15	30	149995	22358	172353
41	16	15	31	196668	18582	215250
49	17	23	40	177521	23810	201331
57	16	27	43	475713	31864	507577
65	17	15	32	307414	29136	336550
72	19	19	38	1378448	30262	1408710
73	16	17	33	236960	35417	272377
86	17	16	33	299536	23646	323182
88	17	16	33	305910	21281	327191
89	15	17	32	331250	34811	366061
97	16	17	33	349200	22990	372190
104	14	16	30	369154	22412	391566
105	20	19	39	391805	18491	410296
113	20	16	36	976871	43977	1020848
120	24	23	47	450947	25980	476927
121	18	25	43	595526	26526	622052
129	20	17	37	759813	23930	783743
136	26	17	43	515506	26171	541677
137	17	19	36	623842	29859	653701
145	17	17	34	595119	26299	621418
152	21	18	39	738931	23263	762194
155	18	19	37	1381600	20194	1401794
161	18	26	44	778209	20636	798845
169	23	18	41	577098	22735	599833
177	17	17	34	49737	27992	77729
184	24	19	43	663831	19800	683631

Table (3) is the summarized of Table (2), and the figure(7), where  $L_i$  denoted to length of text,

Table 3, Time of Encryption,  
Decryption and Total

<b>L<sub>i</sub></b>	<b>TE</b>	<b>TD</b>	<b>TS</b>
001-040	37.00	36.00	73.00
041-056	16.50	19.00	35.5
057-072	17.33	20.33	37.66
073-088	16.67	16.33	33.00
089-104	15.00	16.67	31.67
105-120	21.33	19.33	40.66
121-136	22.00	19.67	41.67
137-152	18.33	18.00	36.33
153-168	18.00	22.50	40.50
169-184	21.33	18.00	39.33

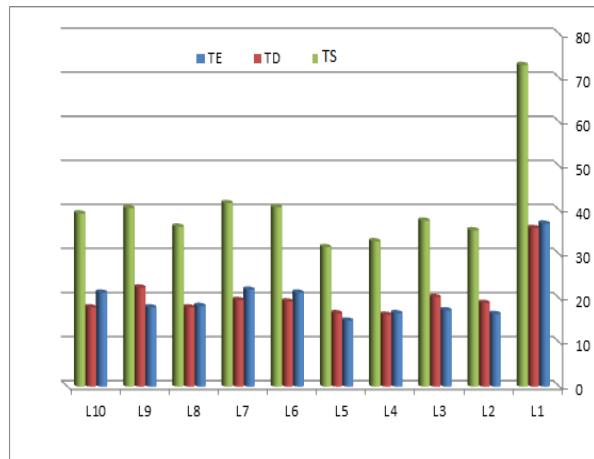


Figure 7, Time of Encryption and Decryption

Table(4) shown the Efficiency Scalar of Encryption(EEE) and Decryption(EED) of the proposed method, depending on time are approximately is equivalent.

Table 4, Efficiency of HVR algorithm

<b>L<sub>i</sub></b>	Time			Efficiency	
	<b>TE</b>	<b>TD</b>	<b>TT</b>	<b>EEE</b>	<b>EED</b>
001-040	37.00	36.00	73.00	4E-04	3E-04
041-056	16.50	19.00	35.5	2E-04	1E-04
057-072	17.33	20.33	37.66	2E-04	2E-04
073-088	16.67	16.33	33.00	2E-04	1E-04
089-104	15.00	16.67	31.67	2E-04	1E-04
105-120	21.33	19.33	40.66	1E-04	1E-04
121-136	22.00	19.67	41.67	1E-04	1E-04
137-152	18.33	18.00	36.33	1E-04	1E-04
153-168	18.00	22.50	40.50	1E-04	9E-05
169-184	21.33	18.00	39.33	1E-04	1E-04

From table 5, note that the fitting ratio(equation (1)) is approximately (100%) between cyphering and deciphering. Where the ratio calculated between the input text( $T: t_i \in T, \forall i$ ), and the resulting text ( $\tilde{T}: \tilde{t}_i \in \tilde{T}, \forall i$ ) from applying cyphering and deciphering on that input text.

$$fitt = 100 * \left[ 1 - \frac{\sum_{i=1}^N (t_i - \tilde{t}_i)^2}{\sum_{i=1}^N (t_i - mt)^2} \right] \quad (1)$$

where:  $mt = mean(T)$

Table 5, Fitting Ratio Scalar

<b>L<sub>i</sub></b>	Fitting ratio	<b>L<sub>i</sub></b>	Fitting ratio
<b>1-40</b>	<b>100.000%</b>	<b>121 – 136</b>	<b>99.999%</b>
<b>41-56</b>	<b>99.900%</b>	<b>137 – 152</b>	<b>99.998%</b>
<b>57- 72</b>	<b>99.999%</b>	<b>153 – 168</b>	<b>100.000%</b>
<b>73 - 88</b>	<b>99.000%</b>	<b>169 – 184</b>	<b>99.989%</b>
<b>89- 104</b>	<b>100.000%</b>	...	...
<b>105 - 120</b>	<b>99.998%</b>	...	...

## VIII. CONCLUSIONS

- \*. A New Genetic algorithm has been proposed in the process of text encryption(decryption).
- \*. The algorithm has been applied in the encryption(decryption) for texts of different sizes.
- \*. Several measures have been introduced to measure the efficiency of the proposed algorithm, including: Time, Efficiency and Fitting Ratio.
- \*. The Scales were applied between the process encryption(decryption) of texts.
- \*. The algorithm has proven successful in encryption(decryption) by up to 100%.
- \*. The encryption method was hybridized using a proposed genetic algorithm.
- \*. Mathematical concepts were used for the proposed method
- \*. The proposed algorithm can be applied to the mobile rather than to the PC.

## IX. Reference

- [1] Hüseyin Bodur and Resul Kara , "Secure SMS Encryption Using RSA Encryption Algorithm on Android Message Application ", Conference: 3RD International Symposium on Innovative Technologies in Engineering and Science, At Valencia, Conference Paper · June 2015.
- [2] Tita Karlita and others, "HILLMAIL: A SECURE EMAIL SYSTEM FOR ANDROID-BASED MOBILE PHONE USING HILL CIPHER ALGORIT", Jurnal Ilmiah KURSOR, Vol.8, No.3, 2016.
- [3] Darshana Patil and Prof. P. M. Chawan , "A Survey on an Enhanced Cryptographic Technique for Messages Encryption and Decryption", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 5, Issue3, 2017.
- [4] Maryam Inayat and others," A Survey on the Encryption Algorithms Used in Cell Phone Messages", International Journal of Emerging Technology and Advanced Engineering, Volume 6, Issue 2, 2017.
- [5] Na Qi Jink Pan Qun Ding,"The Implementation of FPGA-based RSA Public-Key Application and Its Application in Mobile -Phone SMS Encryption System", IEEE International Conference on Instrumentation, Measurement, Computer, Communication and Control, 2011.
- [6] Hemlata Agrawal and Narendra Kahtr , "Image Encryption using Various Transforms-A Brief Comparative Analysis", IEEE, International Conference on Magnetics, Machines & Drives, 2014.
- [7] Al-Hamami, Alaa Hussein and Al-Ani, Saad Abdulaziz, "Information Security Technology and Protection Systems", First Edition, Dar Wael Publishing, Amman, 2007.
- [8] Mohammed Sheikho Mamo, "Principles of Mobile Application Design," Dar Shuaa Publishing & Publishing, Syria [7], 2005.

## AUTHORS PROFILE

Dr. Shatha Abdullah M.Ramadhan is currently a lecturer at Mosul University, College of Computer Sciences and Mathematics/ Software Engineering Department. She received B.Sc. degree in math. science/ University of Mosul in 1997, M.Sc. degree/ University of Mosul in 2000, and Ph.D. degree/ University of Mosul in 2004,. Her research interests are in artificial intelligence, pattern recognition, optimization, information security and image processing,

Hiba Muneer Yahya is currently a assistant lecturer at Mosul University, College of Computer Sciences and Mathematics/ Software Engineering Department.

Ruah Mouad Alyas is currently a assistant lecturer at Mosul University, College of Computer Sciences and Mathematics/ Software Engineering Department.

# Analyzing the Impacts of Solar PV Generation Integration on Radial Distribution Feeder

**Mohsin A. Mari**

Department of Electrical Engineering  
Mehran UET SZAB Campus Khairpur Mir's  
E.mail: mohsinali@muetskhp.edu.pk

**M. Zubair Bhayo**

Department of Electrical Technology  
Benazir Bhutto Shaheed University of Technology and Skill  
Development Khairpur Mir's  
E.mail: muhammadzubairbhayo@gmail.com

**Anwar A. Sahito**

Department of Electrical Engineering,  
Mehran UET, Jamshoro  
E.mail: anwar.sahito@gmail.com

**Basheer A. Kalwar**

Department of Electrical Engineering  
Mehran UET SZAB Campus Khairpur Mir's  
E.mail: basheerahmed@muetskhp.edu.pk

**Abstract**—Human population of the world and its Electrical power demand is increasing day by day. The available fossil fuel energy resources are being depleted day by day. So it is a wise decision to absorb the natural renewable energy resources. Among the other natural resources, solar energy is also a precious available energy source. In Pakistan abundance solar energy can be easily extracted.

In this research work, impacts of solar generation system are analyzed while integrated with 11kV radial distribution feeder. PV system is integrated with feeder in three different ways by using SINCAL software and its impacts in terms of the power loss, voltage profile and short circuit level are analyzed. When PV system is integrated with HT side it results negligible increment in voltage, no change in LT losses, negligible decrement in HT losses and no change in short circuit level. When PV system is connected with LT bus-bar of each transformer, there is significant increment in voltage, small decrement in LT losses, significant decrement in HT losses and smaller increment in short circuit level. When PV system is connected with each load, there is significant increment in voltage, large decrement in LT losses, significant decrement in HT losses and smaller increment in short circuit level.

**Keywords:** Solar PV system; voltage; short circuit level; power losses; distribution feeder.

## 1. INTRODUCTION

Like other developing countries, Pakistan is also facing a critical energy shortage crisis, due to rapid population growth and unsatisfactory available resources[11-17]. The situation is further worsened because of technical and non-technical losses occurring in the existing system[6]. Presently the main resources of power generation in Pakistan are fossil fuels including oil and gas which have no guaranty to be continued for future increased demands. The other major factor of energy wastage is the large distance between generation and utilization. The generation plants are far away from the loads, so the power is carried out through long transmission lines with high voltages. Power reaches to the consumers travelling through primary transmission 500 KV or 220KV, secondary transmission 132 KV or 66 KV then primary distribution 11 KV and secondary distribution 380 V or 230 V and finally through service mains. This whole network offers

various problems and losses to the Electrical Power [10].

If the distance between generation and distribution is reduced by certain means, the power losses will be greatly reduced and we can save our power [5]. This will not only reduce the gap between generation and demand but in future we can fulfill our requirement. The best alternate to this scenario is to install and integrate the distribution generations locally near the end consumers[9]. One of the simple and cheap sources of the distribution generation is the solar power. If the smaller solar PV arrays are designed and connected in series and parallel combination they can generate the required power at required DC voltage which is then converted into AC voltage using inverters[3]. This system may feed the required electrical energy to a home or town locally owing to purchase the costly energy from the WAPDA. The second great benefit is that the consumer's requirement from national grid will be reduced. So when power will not travel through long distances, the power losses will be greatly reduced [8].

Besides the many advantages of renewable solar energy there are some positive and negative impacts of such systems when integrated with the existing working system [7]. These impacts depend upon the size of the distribution generation, techniques of integration, location of integration and the design of the existing system. The main parameters which will be affected are; voltage of the system, power losses and short circuit level of the feeder. To analyze these impacts an existing 11 KV feeder of HESCO named Sachal feeder is taken into consideration and it is simulated using the SINCAL software. First of all normal feeder is simulated and its load flow and short circuit calculations are taken without any solar PV integration. Then three different cases are simulated. In first case three solar generators each of equal to 1 MW rating are installed at three equally spaced locations. In second case one solar generator of half

of the total load capacity of transformer is integrated with the LT bus bar of the transformer. In third case one solar generator of half of the load demand is connected with each load. Finally the results are compared for voltage, power losses and short circuit current level.

Section 2 discusses the solar technology as a distribution generation. Selected feeder and its detail are given in section3. Simulation results are discussed in section 4. Impacts of solar PV integration are given in tables and graphs in terms of the power loss and short circuit level. Finally the conclusion of paper is given in section 5.

## 2. SOLAR PHOTO VOLTAIC GENERATION AND INTEGRATION

In 1839 a French Physist named Becquerel discovered the photo voltaic effect. Up to 1954[7] it remained the constrained of the laboratories. Then Bell laboratories produced the first primary silicon cell which was the quickly improved and used in the united country space program. Solar photovoltaic cells convert light energy directly into Electrical Energy. The generated output voltage of a cell is very small so to increase this voltage, these cells are fabricated in the shape of an array. Each array is called a solar panel and can generate a voltage of 12 V or 24 V DC[4]. To increase this voltage further, these arrays are connected in series with each other, because in series voltage is added. To increase their power rating, these arrays are further connected in parallel. Each small array can provide a power of 35 W, 50 W, 100 W and 150 W. Since this energy totally relies on the presence of sun so naturally variation will occur in the output of these cells. In order to compensate this variation DC Voltage can be regulated by DC-DC converters. If the load is DC it can be feed directly from the output of DC-DC converters, or we can use the batteries which are charged during the day time and we can run our DC load at night time as well from these batteries. To feed the AC loads, this generated DC voltage can be converted into AC voltage by DC-AC converters which are also called inverters. Local AC loads are directly feed from the output of these inverters, and to integrate this AC voltage into an existing AC system transformer is used to step this voltage up according to the voltage available in the system [2].

Solar PV generators can be designed in plenty of the sizes and used for commercial, residential and irrigation load demands. They provide real power to the loads. They are environment friendly and easy to install. They provide a backup support and increase the reliability when integrated with an existing system. They can help the main grid during the peak

hours and provide power to consumers when the load on a particular feeder is beyond its capacity[1]. Large scale solar power generation and integration at either low voltage or high voltage is an emerging trend now a day. Presently power systems of Pakistan are fully overloaded and insufficient for the demand so the solar generation is a precious alternate to overcome the increased load demands. Fig. 1shows a solar power system in which solar panels are arranged in series and parallel combination to fulfill the required voltage and power. There is computer controlled system to monitor the generation. Inverters are shown in a box for converting DC voltage to AC voltage which is then feed to a transformer to step it up. Finally the output from the transformer is connected to the feeder or national grid through three phase distribution system.

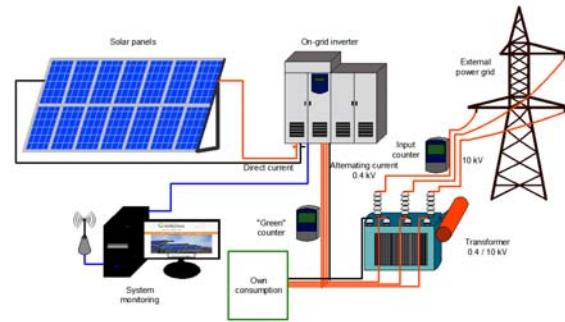


Fig. 1.Solar power plant.

Different countries are offering financial incentives to their residential consumers to install solar power systems to fulfill their local load demands. When a PV solar generation system is integrated with distribution system the power flow and impedance matrices of the system are changed depending upon the size and configuration of the system. So it is important to analyze the impacts of this distribution generation on the radial feeder of HESCO.

## 3. SYSTEM DESCRIPTION

In this research work, 11kV Sachal feeder is selected to observe impacts of solar PV system in terms of voltage, power loss and short circuit current level. Around 3000 residential and commercial consumers are supplied through 41 Pole Mounted Transformers (PMT) having total capacity of 7600 kVA. Dog and Rabbit conductors are used for high tension (H.T) network of 7.42 km.

The entire required feeder data is collected from HESCO and then the required model is designed and simulated using PSS SINCAL software. Separate models are developed for all L.T circuits and linked with H.T circuit for simulation purposes. Four

different cases are analyzed using different simulation circuits. The description of all the cases is listed in Table 1.

TABLE.1 CASES FOR SYSTEM SIMULATION

Case #	Description
I	Existing Network without any solar system
II	Three solar PV systems connected with HT feeder
III	41 separate PV systems are connected with transformer LT side
IV	Each load point is connected with separate PV system

#### 4. RESULTS AND DISCUSSIONS

Simulation results for four mentioned cases are analyzed to observe the positive or negative impacts of solar generation on power losses, short circuit level and voltage on radial distribution feeder.

Table. 2 shows comparison of voltages for some H.T buses in system for analyzed four cases. It can be observed that bus voltages have improved with PV integration. This improvement is due to reduction in current flow as loads are supplied by nearby PV source. Maximum improvement in voltages is observed for Case IV as PV generation is connected nearest to load.

TABLE.2 H.T BUS VOLTAGE COMPARISON

Bus	Voltage (kV)			
	Without PV	PV with H.T	PV at transformer secondary	PV with Load
1	10.993	10.993	10.995	10.996
2	10.993	10.994	10.995	10.995
3	10.846	10.845	10.893	10.901
4	10.968	10.970	10.977	10.980
5	10.963	10.968	10.976	10.977
6	10.961	10.966	10.972	10.976
7	10.960	10.964	10.972	10.975
8	10.959	10.963	10.971	10.974
9	10.954	10.959	10.967	10.972
10	10.953	10.958	10.966	10.971
11	10.952	10.958	10.968	10.970
12	10.951	10.955	10.965	10.970
13	10.948	10.954	10.966	10.968
14	10.942	10.951	10.962	10.964
15	10.939	10.946	10.960	10.962

Short circuit level in MVA is the product of fault current and system rated voltage. It is important consideration for system design as circuit breakers are rated according to calculated short circuit level. Short circuit level depends upon the system impedance and configuration. When a solar PV generation system is connected to an existing system, the impedance of system changes which can cause an increase in the short circuit level. This increase in short circuit level will make it necessary to enhance circuit breaker capacities or incorporate current limiting resistors. Table. 3compares short circuit level for some of the H.T buses for four analyzed

cases. It is evident that short circuit level has increased for all buses but increscent is not significant and therefore same circuit breakers are sufficient.

TABLE.3 H.T. BUS SHORT CIRCUIT LEVEL COMPARISON

Bus	Short circuit Level (MVA)			
	Without PV	PV with H.T	PV at transformer secondary	PV with Load
1	992.133	992.139	992.426	992.507
2	864.708	864.728	864.972	865.047
3	622.170	622.235	622.677	622.846
4	558.005	558.107	558.539	558.719
5	529.146	529.154	529.661	529.841
6	492.426	492.443	492.915	493.094
7	345.357	345.378	345.733	345.895
8	373.434	373.512	373.848	373.905
9	319.995	320.016	320.350	320.505
10	275.644	275.696	275.962	276.099
11	264.743	264.859	265.050	265.182
12	258.943	259.012	259.242	259.370
13	250.076	250.125	250.365	250.489
14	249.262	249.413	249.550	249.675
15	165.860	165.984	166.445	166.589

PV integration to power system will result in change in system losses as current flows are changed. Table.4gives comparison of branch power losses for some of the H.T buses and Table.5gives comparison of power losses for L.T circuits of all 41 transformers. Power losses in most of the branches on H.T circuits are slightly reduced. Similarly all L.T circuits have reduction in power losses. Power loss reduction for all branches is observed when PV is integrated nearest to loads.

TABLE.4 H.T. BRANCH POWER LOSS COMPARISON

Line	Power Losses (kW)			
	Without PV	PV with H.T	PV at transformer secondary	PV with Load
L1-4	0.45	0.41	0.23	0.21
L4-8	2.60	2.53	1.32	1.22
L8-9	0.39	0.34	0.20	0.18
L9-10	0.22	0.22	0.11	0.10
L10-11	0.29	0.27	0.14	0.13
L11-12	0.22	0.22	0.11	0.10
L12-15	0.01	0.01	0.00	0.00
L1-1/1/3	5.01	5.00	2.61	2.21
L1/1/3-1/1/4	1.28	1.25	0.66	0.57
L1/1/4-1/1/4/15	16.64	16.58	8.18	7.66
L1/1/4-1/1/5	0.04	0.04	0.03	0.01
L1/1/5-1/1/6	0.05	0.04	0.03	0.02
L1/1/6-1/1/9	0.20	0.20	0.13	0.08
L12/5-12/7	0.05	0.04	0.02	0.02

TABLE.5.L.T.CIRCUIT POWER LOSS COMPARISON

Transformer	Power Losses (kW)			
	Without PV	PV with H.T	PV at transformer secondary	PV with Load
T1	0.065	0.065	0.064	0.031
T2	0.410	0.405	0.402	0.162
T3	0.138	0.137	0.131	0.06
T4	0.071	0.0705	0.068	0.037
T5	0.378	0.377	0.376	0.159
T6	0.196	0.195	0.193	0.085
T7	0.421	0.420	0.415	0.189
T8	0.033	0.033	0.032	0.033
T9	0.091	0.090	0.089	0.042
T10	0.022	0.020	0.018	0.015
T11	0.066	0.065	0.060	0.031
T12	15.931	15.81	14.494	15.254
T13	0.233	0.230	0.228	0.097
T14	0.237	0.234	0.232	0.1
T15	0.458	0.454	0.447	0.179
T16	0.401	0.398	0.393	0.166
T17	0.014	0.013	0.010	0.005
T18	0.135	0.132	0.130	0.054
T19	0.165	0.163	0.161	0.066
T20	0.027	0.026	0.021	0.017
T21	0.032	0.030	0.029	0.011
T22	0.035	0.033	0.031	0.011
T23	0.044	0.042	0.041	0.016
T24	0.027	0.026	0.021	0.011
T25	0.044	0.040	0.035	0.016
T26	0.044	0.041	0.039	0.016
T27	0.049	0.044	0.040	0.016
T28	0.049	0.042	0.031	0.015
T29	0.013	0.013	0.012	0.005
T30	0.311	0.310	0.308	0.144
T31	0.336	0.331	0.325	0.154
T32	0.061	0.060	0.059	0.029
T33	0.144	0.141	0.138	0.064
T34	0.028	0.025	0.021	0.012
T35	0.036	0.031	0.026	0.015
T36	0.143	0.138	0.131	0.075
T37	0.068	0.064	0.059	0.029
T38	0.134	0.127	0.117	0.059
T39	1.124	1.119	1.095	0.514
T40	0.026	0.021	0.018	0.009
T41	0.199	0.194	0.181	0.094

Fig. 2 shows graphical comparison of total power losses of the Sachal feeder for four analyzed cases. It is observed that there is considerable power loss reduction with PV integration to selected feeder. Maximum power loss reduction is achieved if PV generation is connected near the loads.



Fig. 2.Solar power performance

## 5. CONCLUSIONS

Solar PV generation is increasing rapidly around the world. Large PV power generation facilities are being installed by utilities. Quaid e Azam solar park is one of the first large solar generations in Pakistan. Small PV power generation facilities are also being installed around the country.

In this research work, effects of solar PV generation on distribution network are analyzed using simulation on PSS-SINCAL software. 11 kV Sachal feeder is modelled and simulated after collection of real time data. Four different simulations are performed to observe effects of different locations

- Without any PV
- PV distributed on H.T network
- PV connected with transformer bus on L.T
- PV connected with loads representing individual PV generation by consumers

Comparison of results for four simulation cases shows that:

H.T and L.T power losses are reduced with PV integration

- Highest improvement is observed when PV generation is connected with loads
- Lowest is observed with PVs connected on H.T network
- Power and current flows have changed with PV integration

It is therefore concluded that PV integration will improve system performance specially those small PV generations installed by consumers.

## ACKNOWLEDGEMENT

Authors are thankful to Mehran University of Engineering and Technology Jamshoro for providing necessary resources as well as authors are thankful to Dr Mazhar Hussain Baloch assistant professor for his kind support regarding review of this paper.

## REFERENCES

- [1].Adnan, S., A.H. Khan, S.Haider, and R. Mahmood, (2012), "Solar Energy Potential in Pakistan" *Journal of Renewable and Sustainable Energy*, 4(3), 327-401
- [2].Kyritsis, A., Voglitsis, D., Papanikolaou, N., Tselepis, S., Christodoulou, C., Gonos, I., &Kalogirou, S. A. (2017). "Evolution of PV systems in Greece and review of applicable solutions for higher penetration levels". *Renewable Energy*, 109, 487-499.
- [3].Ravindran, V. and V.Aravinthan, (2013), "Feeder Level Power Loss Reduction through Reactive Power Control with Presence of

Distributed Generation." *IEEE Power & Energy Society General Meeting, Canada, 1-5.*

[4].Roediger, S.; R. Yan, and T. Saha, (2012), "Investigation of the Impacts of Three-Phase Photovoltaic Systems on Three-Phase Unbalanced Networks", in *IEEE PES General Meeting, California, USA, 1-8.*

[5].Sahito, A.A., I.A. Halepoto, and S.M. Tunio (2014), "Application of Infrared Thermography in Power Distribution System", *Mehran University Research Journal of Engineering & Technology*, 33(3), 352-358.

[6].Sahito, A.A., Z.A. Memon, G.B.Buriro, S.A.Memon, M.J. Jumani (2016)"Voltage Profile Improvement of Radial Distribution Feeder through Distributed Generation", *Sindh University Research Journal (Science Series)*, 48(3), 497-500.

[7].Shami, S.H., J. Ahmad, R.Zafar, M.Haris, and S. Bashir (2016), "Evaluating Wind Energy Potential in Pakistan's Three Provinces, with Proposal for Integration into National Power Grid." *Renewable and Sustainable Energy Reviews*, 53, 408-421

[8].Siddique, S., and R. Wazir, (2016), "A Review of the Wind Power Developments in Pakistan." *Renewable and Sustainable Energy Reviews*, 57, 351-361.

[9].Umer, F.,N. K.Afridi,A.A. Sahito, and A.Z. Pathan, (2013), "Critical Analysis of Capacitors as a Potential Solution to Achieve Optimum Level Voltage Regulation in HESCO Network." *IEEE International Conference on Technological Advances in Electrical, Electronics and Computer Engineering (TAECEC), Turkey*, 488-497.

[10].Uqaili, M. A., A.A. Sahito, I.A. Halepoto, Z.A.Memon,, and S. B.Dars, (2014), "Impact of Distributed Generation on Network Short Circuit Level", Proceedings of *IEEE 4th International Conference on Wireless Communications, Vehicular Technology, Information Theory and Aerospace &Electronic Systems (VITAE), Denmark, 1-5.*

[11] Baloch, Mazhar H., Ghulam S. Kaloi, and Zubair A. Memon. "Current scenario of the wind energy in Pakistan challenges and future perspectives: A case study." *Energy Reports* 2 (2016): 201-210.

[12] Baloch, Mazhar Hussain, Ghulam Sarwar Kaloi, and Jie Wang. "Feasible Wind Power Potential from Costal Line of Sindh Pakistan." *Research Journal of Applied Sciences, Engineering and Technology* 10.4 (2015): 393-400.

[13] Baloch, Mazhar H., et al. "A research on electricity generation from wind corridors of Pakistan (two provinces): a technical proposal for remote zones." *Sustainability* 9.9 (2017): 1611.

[14] Kaloi, Ghulam Sarwar, et al. "Wind Energy Potential at Badin and Pasni Costal Line of Pakistan." *International Journal of Renewable Energy Development* 6.2 (2017): 103.

[15] Kaloi, Ghulam Sarwar, Jie Wang, and Mazhar Hussain Baloch. "Active and reactive power control of the doubly fed induction generator based on wind energy conversion system." *Energy Reports* 2 (2016): 194-200.

[16] Baloch, Mazhar H., Jie Wang, and Ghulam S. Kaloi. "Stability and nonlinear controller analysis of wind energy conversion system with random wind speed." *International Journal of Electrical Power & Energy Systems* 79 (2016): 75-83.

[17] Zhu, Zhijian, Jie Wang, and Mazhar H. Baloch. "Dynamic economic emission dispatch using modified NSGA-II." *International Transactions on Electrical Energy Systems* 26.12 (2016): 2684-2698.

# Which NoSQL Database to Combine with Spark for Real Time Big Data Analytics ?

Omar HAJOUI, Mohamed TALEA  
LTI Laboratory, Faculty of Science Ben M'Sik  
Hassan II University, Casablanca, Morocco  
[{hajouio,taleamohamed}@yahoo.fr](mailto:{hajouio,taleamohamed}@yahoo.fr)

**Abstract— Big Data is an evolution of Business Intelligence (BI).** Whereas traditional BI relies on data warehouses limited in size (some terabytes) and it hardly manages unstructured data and real-time analysis, the era of Big Data opens up a new technological period offering advanced architectures and infrastructures allowing sophisticated analyzes taking into account these new data integrated into the ecosystem of the business . In this article, we will present the results of an experimental study on the performance of the best framework of Big Analytics (Spark) with the most popular databases of NoSQL MongoDB and Hadoop. The objective of this study is to determine the software combination that allows sophisticated analysis in real time.

**Keywords-** *big data analytics; NoSQL databases; Apache Spark ; Hadoop; MongoDB, performance .*

## I. INTRODUCTION

The Big Data phenomenon, for companies, covers two realities: on the one hand this explosion of data continuously, on the other hand the capacity to process and analyze this great mass of data to make a profit. With Big Data, organizations can now manage and process massive data to extract value, decide and act in real time.

NoSQL databases were developed to provide a set of new data management features while overcoming some limitations of currently used relational databases [1]. NoSQL databases are not relational and they don't require a model or structure for data storage, which facilitates the storage and data search. In addition, they allow horizontal scalability, it gives administrators the ability of increasing the number of server machines to minimize overall system load. The new nodes are integrated and operated in an automatic manner by the system. Horizontal scalability reduces the response time of queries with a low cost.

In relation to the NoSQL databases (Hadoop, MongoDB, Cassandra, Hbase, Radis, Riak...., etc.), a new profession appeared "the data scientist". Data science is the extraction of knowledge from data sets [2, 3]. It employs techniques and theories derived from several other broader areas of mathematics, mainly statistics, probabilistic models, machine learning. Thus, to develop algorithms in a distributed environment, the analyst must master tools of big data analytics (Mahout, MapReduce, Spark and Storm) and learn the syntax of functional languages to use Scala, Erlang or Clojure.

Big data analytics therefore favors a return to grace of functional languages and robust methods: decision tree [4, 5], and random forest [6], k-means [7], Naive Bayes classifier [8], easily distributable (MapReduce) on thousands of nodes.

For collected data storage, any NoSQL database can fulfill this role. However, the need to analyze this data pushes us to choose this database carefully. Especially in the field of Big Data, the analytic part becomes more and more important. For advanced, real-time analytics, the best framework you can use is Apache Spark [9, 10]. According to the official version, Spark uses the hadoop HDFS file system.

In a previous study [11] based on a multicriteria analysis method, the MongoDB system obtained the highest score. Today, this result was confirmed. This system has become popular [12]. According to a white paper [13] published by MongoDB, The combination of the fastest analysis engine (Spark) with the fastest-growing database (MongoDB) allows companies to easily perform reliable real-time analysis. This led us to compare Spark's performance against the most popular NoSQL databases, MongoDB and Hadoop. In this article, we will present and discuss the results of our experimental study. Thus, we will determine the software combination that allows giving sophisticated analyzes in real time.

This paper is organized as follows: Section II presents big data analytics on Hadoop and MongoDB. In section III, we present the results of an experimental study on the performance of the framework Spark with MongoDB and Hadoop. Section IV provides a conclusion.

## II. BIG DATA ANALYTICS

In this part, we will introduce the data analysis technologies used on Hadoop and MongoDB.

### A. Big Data Analytics on Hadoop

The first integrated solution with Hadoop for data analysis is the MapReduce framework. MapReduce is not in itself an element of databases. This distributed information processing approach takes an input list, produces one in return.it can be used for many situations; it is well suited for distributed processing needs and decision-making processes.

MapReduce defined in 2004 in an article written by Google. The principle is simple: to distribute a treatment, Google imagined a two-step operation. First, an assignment of operations on each machine (Map) followed processing by a grouping of results (Reduce). The needs of Google that gave birth to MapReduce are twofold: how to handle gigantic volumes of unstructured data (web pages to analyze to feed the Google search engine, or the analysis of the logs produced by the work of its indexing engines, for example), to derive results from calculations, aggregates, summaries ... in short, from the analysis.

The free reference implementation of MapReduce is called Hadoop, a system developed by a team led by Doug Cutting, in Java, for the purposes of its Nutch distributed indexing engine for Yahoo! Hadoop directly implements the Google document on MapReduce, and bases its distributed storage on HDFS (Hadoop File System), which implements the Google document on GFS (Google File System). Then, the Hadoop MapReduce Framework (YARN) implemented by several NoSQL databases such as Hbase, Cassandra, etc.

Then, Facebook developed the HQL language (Hive language query) on Hive. Close to SQL to query HDFS. Another language, called Pig, developed by Yahoo similar in its syntax to Perl and aimed at the goals of Hive. In addition, cloudera, another Hadoop distribution, integrates Impala's queries engine. Analysts and data scientists to perform analysis on data stored in Hadoop via SQL tools or business intelligence tools favor this latest one. The Mahout project provides algorithms implementations for business intelligence. It provides, for example, machine-learning algorithms (Kmeans, Random Forest).

#### B. Big Data Analytics on MongoDB

MongoDB is an open-source document-oriented database designed for exceptionally high performance and developed in C++. Data is stored and queried in BSON format similar to JSON. It has dynamic and flexible schemas, making data integration easier and faster than traditional databases. Unlike NoSQL databases that offer basic queries. Developers can use MongoDB native queries and data mining capabilities to generate many classes of analysis, before having to adopt dedicated frameworks such as Spark or MapReduce for more specialized tasks.

Several organizations including McAfee, Salesforce, Buzzfeed, Amadeus, KPMG and many others rely on MongoDB's powerful query language, aggregations and indexing to generate real-time analytics directly on their operational data. MongoDB users have access to a wide range of queries, projection and update operators that support real-time analytic queries on operational data:

- **The MongoDB Aggregation Pipeline** is similar in concept to the SQL GROUP BY statement, enabling users to generate aggregations of values returned by the query (e.g., count, minimum, maximum, average, intersections) that can be used to power analytics dashboards and visualizations.

- **Range queries** returning results based on values defined as inequalities (e.g., greater than, less than or equal to, between)
- **Search queries** return results in relevance order and in faceted groups, based on text arguments using Boolean operators (e.g., AND, OR, NOT), and through bucketing, grouping and counting of query results.
- MongoDB provides native support for **MapReduce**, allowing complex JavaScript processing. Multiple MapReduce jobs can run simultaneously on the same server and on fragmented collections.
- **JOINS , Graph queries , Key-value queries ...**

#### C. Big Data Analytics on Hadoop

The MapReduce framework, despite being widely used by companies for the analysis of Big Data, the response time is not satisfactory and its programs executed only in the form of a batch. After a map or reduce operation, the result must be written to disk. This disk-written data allows mappers and reducers to communicate with each other. It is also the write on disk, which allows a certain tolerance to the failures: if a map or reduce operation fails, it is enough to read the data from the disk to take again, where we were. However, these writings and readings are time consuming. In addition, the expression set composed exclusively of map and reduce operations is very limited and not very expressive. In other words, it is difficult to express complex operations using only this set of two operations.

Apache Spark is an alternative to Hadoop MapReduce for distributed computing that aims to solve both of these problems. The fundamental difference between Hadoop MapReduce and Spark is that Spark writes data in RAM, not on disk. This has several important consequences on the speed of calculation processing as well as on the overall architecture of Spark.

Spark offers a complete and unified framework (Figure 1) to meet the needs of Big Data processing for various datasets, various by their nature (text, graph, etc.) as well as by the type of source (batch or real-time flow). It allows to quickly write applications in Java, Scala or Python and includes a set of more than 80 high-level operators, it is possible to use it interactively to query the data from a shell, in addition to the operations of Map and Reduce, Spark supports SQL queries and data streaming and offers machine learning and graph-oriented processing functions. Developers can use these possibilities in stand-alone or by combining them into a complex processing chain.

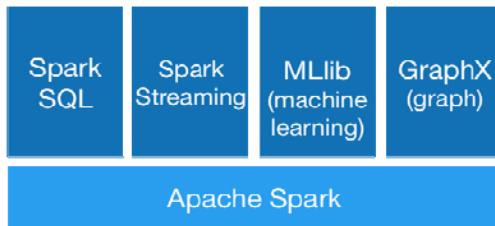


Figure 1: Apache Spark Ecosystem

Spark's programming model is similar to MapReduce, except that Spark introduces a new abstraction called Resilient Distributed Datasets (RDDs). Using RDDs, Spark can provide solutions for several applications that previously require the integration of multiple technologies, including SQL, streaming, machine learning and graph processing.

A Dataset is a distributed collection of data. It can be viewed as a conceptual evolution of RDDs (Resilient Distributed Datasets), historically the first distributed data structure used by Spark. A DataFrame is a Dataset organized into columns that have names, such as tables in a database. With the Scala programming interface, the DataFrame type is simply the alias of the Dataset [Row] type.

It is possible to apply **actions** to the Datasets, which produce values, and **transformations**, which produce new Datasets, as well as certain functions that do not fit into either category.

```

//For example to create a Dataset from the text file named LICENSE:
scala> val texteLicence = spark.read.textFile("LICENSE")

//An example of action:
scala> texteLicence.count() // number of lines

/* We can use a transformation to build a Dataset containing only
the lines of texteLicence which contain "Copyright", to return a
table with its 2 first lines:*/
scala> val lignesAvecCopyright = texteLicence.filter(line =>
line.contains("Copyright"))

scala> lignesAvecCopyright.take(2)
  
```

Figure 2: Spark Command lines Example

Spark exposes RDDs through a functional programming API in Scala, Java, Python, and R, where users can simply pass local functions to run on the cluster.

### III. COMPARISON

#### A. The Experiments Results

We made the comparison on files of the same size and type (.CSV).The test files are available on this link "<https://catalog.data.gov/dataset/crimes-2001-to-present-398a4>". We copied each file to the Hadoop file system. Then the same file imported by MongoDB.

We did the test on one node, three nodes and four nodes. The machines used having the following configuration:

- 8GB RAM
- Linux Fedora 26
- 120 GB SSD
- 6th generation i5 processor

Table 1: Spark's performance with Hadoop and MongoDB

Nodes	File size (GB)	Action	Hadoop	MongoDB
1	1,55	first	96 ms	77 ms
		count	10 s	2,0 min
3	3,11	first	90 ms	65 ms
		count	19 s	3,4 min
4	4,66	first	0,1 s	57 s
		count	29 s	5,3 min

These results are illustrated in the following figure:

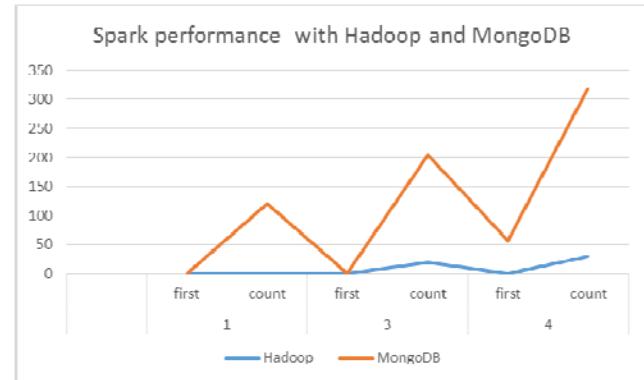


Figure 3: Comparison of Spark's performance versus Hadoop and MongoDB

#### B. Results Interpretation

According to the results of this study, the execution time of the first operation that looks for the first record of the file is the same on Hadoop or MongoDB, sometimes Spark is faster with MongoDB, but the execution time of the operation count that requires the change of the entire file in memory in a RDD, Spark is much faster with Hadoop.

For the moment, Hadoop remains the best global storage solution with administration that is more advanced, security and monitoring tools. This choice, Oracle did for its brand new data discovery and analysis solution, Big Data Discovery. The product installs on a Hadoop cluster (exclusively Cloudera) and relies heavily on Spark for its treatments.

#### IV. CONCLUSION

In this article, we presented the results of an experimental study on the performance of the best framework of Big Analytics (Spark) with the most popular databases of NoSQL MongoDB and Hadoop. The aim of this study is to determine the software combination that allows sophisticated analysis in real time. According to the results of this study, Spark is much faster with Hadoop.

#### REFERENCES

- [1] NoSQL , <http://nosql-database.org/>, 2018
- [2] Vasant Dhar, « Data Science and Prediction », Communications of the ACM, no 12, décembre 2013, p. 64-73
- [3] Davenport et DJ Patil « Data Scientist: The Sexiest Job of the 21st Century », Harvard Business Review, 2012
- [4] Breiman, L., Friedman, J., Olshen, R., and Stone, C. (1984). Classification and Regression Trees. Chapman & Hall, New York.
- [5] L. Breiman. Bagging predictors. Machine Learning, 24(2), 1996.
- [6] L. Breiman. Random forests. Machine Learning, 45, 2001.
- [7] MacQueen J. B. (1967). Some Methods for classification and Analysis of Multivariate Observations. Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability 1. University of California Press. pp. 281-297.
- [8] Marron M. E. (1961). Automatic Indexing : An Experimental Inquiry, Journal of the ACM (JACM), VOL. 8 : Iss. 3, pp 404-417.
- [9] Zaharia, M., Xin, R. S., Wendell, P., Das, T., Armbrust, M., Dave, A., ... & Ghodsi, A. (2016). Apache Spark: A unified engine for big data processing. Communications of the ACM, 59(11), 56-65.
- [10] Gopalani, S., & Arora, R. (2015). Comparing apache spark and map reduce with performance analysis using K-means. International Journal of Computer Applications, 113(1).
- [11] Omar, H., Rachid, D., Mohammed, T., Zouhair, I., B.: An Advanced Comparative Study of the Most Promising NoSQL and NewSQL Databases With a Multi-Criteria Analysis Method. Journal of theoretical and applied information technology, Vol81, No3.
- [12] Solid IT (2018) , <https://db-engines.com>
- [13] <https://www.mongodb.com/collateral/apache-spark-and-mongodb-turning-analytics-into-real-time-action>

# Scheduling problems from workshop to collaborative mobile computing: A state of the art

Ghassan Fadlallah, Djamal Rebaine, Hamid Mcheick

Département d'informatique et de mathématique, University of Quebec in Chicoutimi

555, boulevard de l'Université, Saguenay, Québec, Canada G7H 2B1

[Ghassan.Fdalallah1@uqac.ca](mailto:Ghassan.Fdalallah1@uqac.ca) [Djamal\\_Rebaine@uqac.ca](mailto:Djamal_Rebaine@uqac.ca) [Hamid\\_Mcheick@uqac.ca](mailto:Hamid_Mcheick@uqac.ca)

**Abstract-** This survey reviews the latest literature related to scheduling problems which is closely related to load balancing problems. It is noted that they are often used with the same meaning. In fact, it is not efficient to use one without the other. This is because the scheduling problem is to determine the order of tasks execution on available devices, while load balancing seeks to balance these tasks between these devices. The motivation of this work comes from the need to have, in one paper, a comprehensive idea of these problems with an in-depth view of the involved research tendencies. Several scheduling schemes under different constraints and optimization criteria are discussed. We observed that the rapid technological development at the level of machinery and equipment is accompanied by intensive use of these devices. This requires the enhancement and improvement of scheduling algorithms and the tendency is more and more towards the heuristic and approximate algorithms. As the scheduling schemes range from workshops to Cloud, Fog and Edge computing segments of the collaborative mobile computing, we argue that they have not yet been used effectively in its third segment: individual mobile networks. These networks can play the most effective role, in catastrophic situations, to overcome the problem of telephony/internet communication traffic with the cheapest or free cost. We aim to motivate research on scheduling issues to this segment of collaborative mobile computing that becomes indispensable in urgent these cases as: Oregon, floods, earthquake, terrorist attacks, etc., when almost everything is damaged or not accessible except our small mobile devices and ubiquitous resources.

## 1 INTRODUCTION

Accelerated technological development constantly imposes real challenges in various fields, especially in mobile networks because of the rapid advances in computer architecture, mobile devices and wireless communications. This has led to the transition from big devices network (desktops and laptops) to small mobile devices connected by high bandwidth wireless. The advances in hardware and software technologies have sparked increased interest in the use of these mobile devices within the large scale parallel and distributed systems in different fields such as databases, defense, real-time and commercial applications. The performance of any system designed to operate a large number of devices depends on the tasks scheduling satisfying the workload distribution across these devices.

Scheduling involves allocating resources over time to perform a collection of tasks [4, 6]. The need of scheduling started first in factories and industries before becoming a de facto technique in multiprocessor computers.

Due to the explosion of mobile and wireless technologies, and despite the fourth generation (and soon the fifth) deployment of wireless communication systems, several challenges still remain to be solved. These challenges include the spectrum crisis, high energy consumption, the ever-increasing demand for high data rates and the mobility required by new wireless applications.

The latest technological advances in this area can be a solution to the intensive use of applications on mobile devices. As this situation has captured users, because it gives them freedom in terms of place of work and time to pursue their jobs and interests. This increases the demand for using these devices. Unfortunately, the rate of this growing demand is still higher than the technological growth. Therefore, this reinforces the trend of improving scheduling and load balancing techniques.

According to Leung and Anderson in [1], a scheduling process involves modeling a range of different environments which differ in the way the information is released. They distinguish paradigms: the static scheduling, when all jobs with related information are available at the start of the horizon, and the dynamic scheduling where jobs have different release or available times. The authors pointed out that the decision maker must optimize (usually minimize) a given objective function. There are different categories of policies classified generally into: a) the class of static policies; when the decision maker has to specify at the outset all actions to be taken during the evolution of the process, b) the class of dynamic policies; decisions are made at any time as a function of all the information that has become available up to that point.

In scheduling paradigm, a distinction is also made between i) offline deterministic scheduling, ii) stochastic scheduling, and iii) online scheduling [1,7,8,9].

*In offline deterministic scheduling*, all information or data with regard to the problem is known *a priori* including: the number of jobs, their release dates, due dates, weights, etc. The resulting problem is known as a combinatorial optimization problem subject to some given constraints.

*In stochastic scheduling* the number of jobs is fixed and known in advance. However, most or all of the parameters describing a job such as processing times, release dates and due dates are considered as random variables from known distributions.

*In online scheduling*, there is even less information known before hand, it is released gradually to the decision maker. The decision maker knows nothing in advance about the release dates or processing times.

To summarize, offline deterministic scheduling deals with perfect information, stochastic scheduling with input that is partially stochastic, while online deterministic scheduling deals with input that is known gradually as it arrives to the system. However, a model which mixes the above models may also be the subject of interesting studies. Indeed, Vredeveld [2] addressed the stochastic online scheduling (SOS) model. In this model, jobs arrive in an online manner and as soon as a job becomes known, the scheduler only learns about the probability distribution of the processing time and not the actual processing time. Both online scheduling and stochastic scheduling are special cases of this model.

Let us note that deterministic scheduling models are based on predictive approaches that do not take into account the presence of disturbances, and evaluate the scheduling solution in terms of estimated data. In practice, however, this sort of scheduling becomes quickly unfeasible and returns poor performance. Indeed, in practice, scheduling environments are usually subject to significant amounts of randomness.

As a result it is not of interest to spend an enormous amount of time figuring out a supposedly optimal solution when within a few hours random events will change the structure of the problem or the list of jobs [5]. So, the hypothesis of determinism of scheduling problems is considered as restrictive, and the problem of scheduling with uncertainty management has been raised and is of interest to several researchers [3]. This has led to motivate research in the dynamic scheduling methods which consist in (re)allocating resources at run time [128] *i.e.* make decisions in real time given the state of resources and the progress of different tasks over time<sup>1</sup>. These methods use approaches other than the predictive ones, which until recently were known as proactive, reactive and hybrid which includes two sub-types: predictive-reactive and proactive-reactive approaches. The proactive approach computes by anticipation a scheduling solution by taking into account *a priori* knowledge about probable uncertainties. The reactive approach, another on-line approach, builds real-time scheduling solutions by taking into account any kind of uncertainty that may arise. Finally, it is possible to combine on-line (reactive) and off-line (proactive) approaches in order to get the advantages offered by the two models [17].

### 1.1 Definitions and notations in scheduling problems

The characteristics of jobs may be of the following: preemptive or non-preemptive, resumable or non-resumable, independent or linked. The latter is represented usually by a precedence graph. This graph is a directed acyclic graph that specifies the precedence constraints between tasks execution [136]. A job is called non-preemptive if the processing on this job which is assigned to a machine is processed until its completion. On the other hand, if this processing is interrupted before its completion and reassigned to either the same machine or some other machine, that type of job is called preemptive [10].

A job is said to be resumable, if it has been interrupted due to a machine non-availability period, it can be resumed without needing to be restarted after the machine becomes available again. But, the non-resumable job has to be restarted every time it is interrupted [135].

Polynomial time [155]: The time complexity of an algorithm is said to be of polynomial time if the running time of this algorithm is  $O(p(n))$ , where  $p(n)$  is a polynomial and  $n$  is the size of the input of the problem being solved.

The setup time is defined in [20] as the time required to prepare the necessary resource, as machines or people, to perform a task or job operation. The setup cost is the cost to set up any resource used prior to the execution of a task (for more details see e.g. Allahverdi and Soroush [20]).

In the reminder of this paper, we will be using the following notations [127]:  $C_i$ ,  $F_i = C_i - r_i$ ,  $L_i = C_i - d_i$ ,  $w_i$ ,  $T_i = \max(0, L_i)$ ,  $E_i = \{d_i - C_i\}$ ,  $r_i$ ,  $p_i$ ,  $d_i$  are respectively completion time, flow time, lateness, relative weight,

---

<sup>1</sup> Groupe d'Ordonnancement Théorique et Appliquée (GOThA)

tardiness, earliness, release date, processing time, due date of job  $i$ . Table I summarises the most used criteria of performance to evaluate the quality of a scheduling solution.

TABLE I  
THE PERFORMANCE MEASURES TO EVALUATE THE QUALITY OF A SCHEDULING SOLUTION

Measure	Max	Total	Total weighted	Average	Average weighted
Completion time $C_i$	$C_{max} = \max_{1 \leq i \leq n} C_i$	$C_{total} = \sum_{i=1}^n C_i$	$wC_{total} = \sum_{i=1}^n w_i C_i$	$C_{average} = \frac{1}{n} \sum_{i=1}^n C_i$	$wC_{average} = \frac{1}{n} \sum_{i=1}^n w_i C_i$
Flow time $F_i = C_i - r_i$	$F_{max} = \max_{1 \leq i \leq n} F_i$	$F_{total} = \sum_{i=1}^n F_i$	$wF_{average} = \frac{1}{n} \sum_{i=1}^n w_i F_i$	$F_{average} = \frac{1}{n} \sum_{i=1}^n F_i$	$wF_{average} = \frac{1}{n} \sum_{i=1}^n w_i F_i$
Lateness $L_i = C_i - d_i$	$L_{max} = \max_{1 \leq i \leq n} L_i$	$L_{total} = \sum_{i=1}^n L_i$	$wL_{average} = \frac{1}{n} \sum_{i=1}^n w_i L_i$	$L_{average} = \frac{1}{n} \sum_{i=1}^n L_i$	$wL_{average} = \frac{1}{n} \sum_{i=1}^n w_i L_i$
Tardiness $T_i = \max(0, L_i)$	$T_{max} = \max_{1 \leq i \leq n} T_i$	$T_{total} = \sum_{i=1}^n T_i$	$wT_{average} = \frac{1}{n} \sum_{i=1}^n w_i T_i$	$T_{average} = \frac{1}{n} \sum_{i=1}^n T_i$	$wT_{average} = \frac{1}{n} \sum_{i=1}^n w_i T_i$
Earliness $E_i = \{d_i - C_i\}$	$E_{max} = \max_{1 \leq i \leq n} E_i$	$E_{total} = \sum_{i=1}^n E_i$	$wE_{average} = \frac{1}{n} \sum_{i=1}^n w_i E_i$	$E_{average} = \frac{1}{n} \sum_{i=1}^n E_i$	$wE_{average} = \frac{1}{n} \sum_{i=1}^n w_i E_i$

Competitive ratio is the way to evaluate the performance of an online algorithm. The idea is to evaluate the quality of an on-line algorithm compared to an algorithm that receives the complete information. A competitive ratio is defined as [156]:

$$\max_{e \in E} \frac{\text{Cost of executing the plan that does not know } e \text{ in advance}}{\text{Cost of executing the plan that knows } e \text{ in advance}}$$

Scheduling problems may be solved using either meta-heuristic algorithms or heuristic algorithms [157] i.e iterative or constructive methods that may deliver approximate solutions within a reasonable time. The popular ones are as follows:

- Genetic algorithms are initially developed to meet specific needs in biology. In the context of combinatorial optimization applications, an analogy is developed between an individual in a population and a solution of a problem in the global solution space [157].
- The simulated annealing method; used in metallurgy to improve the quality of a solid and seeks a state of minimal energy that corresponds to a stable structure of the solid. The simulated annealing method is designed to solve local minima problems. [158].
- PSO (Particle Swarm Optimization algorithm) is a cooperative, population-based global search swarm intelligent metaheuristic, presented by Kennedy and Eberhart in 1995 [33]. This is a powerful optimization technique for solving multimodal continuous optimization problems [34]. It is also a population based stochastic optimization technique which has become popular due to its effectiveness and low computational cost.
- Longest Processing Time rule (LPT): Jobs with large processing time values are prioritized for scheduling. So, tasks are organized in descending order of their processing times. Shortest Processing Time rule (SPT): Jobs with small processing time values are prioritized for scheduling. So, tasks are organized in ascending order of their processing times [7].

Contingent schedule: “a contingent schedule allows different task resource assignments depending on how the execution of the schedule has proceeded so far. A contingent schedule can be viewed as a tree that assigns a possibly different task resource assignment for every possible execution [117].”

## 1.2 Context of the research topic

The main objective in the machines scheduling theory is to find the best solutions everywhere in that broad scope with various areas, such as production, medical, military, informatics services and telecommunications, etc. This scheme aimed to adopt different classifications of these problems each of which has been named, over the years, by different names among the following: static, deterministic, predictive, offline, stochastic, dynamic, proactive, reactive, online, etc.

In the area of mobile networking, Cloud computing and Fog or Edge computing, which are the recent emerging domains of parallel machines scheduling, it is well justified to focus our study on the multi-machine model, dedicated machines and parallel machine model [7,10].

The parallel machine models are usually classified as follows:

1. Identical parallel machines: all the machines are identical in terms of their speed. Every job will take the same amount of processing time on each of the machines.

2. Uniform (proportional) parallel machines: the machines have different speeds, whereas for each job, its processing times on the machines are inversely proportional to the speeds of those parallel machines.
3. Unrelated parallel machines: the machines have different speeds with arbitrary processing times of jobs. In this type of scheduling, there is no relation amongst the processing times of a job on the parallel machines.

Researches on scheduling problems have expanded over decades. These researches differ according to the perspective from which they are based, since there is no common perspective to address this issue. In fact, many surveys have been published over the last few years treating these problems of several points of view.

This paper is an attempt towards a comprehensive overview of dealing with scheduling problems based on a significant number of surveys and individual articles that have been published in the literature over the past decade.

This survey will focus on scheduling problems with multiple parallel machines or devices. This is motivated by the fact that the problem of scheduling, as mentioned above, has become, according to the accelerated technological progress, an evident requirement in recently emerging fields, such as modern factories, mobile networks, cloud computing, fog computing, smart cities, etc.

The remainder of this paper is organized as follows. Section 2, 3 and 4 present respectively recent surveys and articles on scheduling problems in i) industry, ii) Cloud and Edge computing, and iii) mobile networks. These sections are presented in two forms: the first details some important works, while the second summarizes other surveys in a table illustrating their most important characteristics such as: reference, scheduling area, context, scheduling environment, conclusion, suggestions and tendencies. Section 5 highlights the discussion of the most important ideas, suggestions and tendencies in these surveys. In Section 6 we conclude the examined papers in terms of the recent scheduling trends in emerging areas such as Cloud, Edge / Fog and Mobile Computing.

## 2 SURVEYS AND ARTICLES ON CLASSICAL AND INDUSTRIAL SCHEDULING PROBLEMS

Many research algorithms and surveys have been published in the area of scheduling problems. In this section, we analyse and discuss briefly these works to understand the different approaches and algorithms used for their resolutions.

Saidy *et al.* [11] studied scheduling problems under various constraints such as activity duration, release dates, due dates and precedence constraints, and the availability of resources (or machines) that could affect these problems. In fact, these machines might not be available during certain time periods called holes for convenience [12]. Their study is based on the classification of these problems into two classes: deterministic and stochastic.

In this survey [11], Saidy *et al.* presented machine scheduling problems that have been studied in the literature with availability constraints in the resumable, semi-resumable and non-resumable cases, within different environments: single machine, parallel machines, flow shop, job shop, open shop, flexible flow shop and flexible job shop. They mentioned papers dealing with single machine and also parallel machine problems, having resumable, non-resumable and "crossable" availability constraints. They have defined as "crossable" the unavailability period that allows an operation to be interrupted and resumed after a period of time. Although, an unavailability period that prevents the interruption of any operation, even if the operation is resumable, as "non-crossable".

Saidy *et al.* stated that in stochastic models the following parameters are not known before time: the processing times, the release dates, the starting time and the duration of the unavailability period. However, they assumed that the distributions of the processing times, due dates (deadlines), repair time and time at which breakdown occurs are known at time 0. The uptime and downtime of machines and the jobs processing requirement are assumed to be independent identically distributed random variables.

They pointed out that heuristics are used for problems in both cases: when machines are continuously available and when breakdowns occur. At their days, they considered that this document may be a good reference for those interested in sequencing and job scheduling issues in the context of limited resource availability.

Most of the heuristics with error bound analysis have been gathered in this study [11], noting that these heuristics, in some cases, produce optimal solutions. Also, the known Polynomial (P) and pseudo-Polynomial (pseudo-P) models were summarized in a single table whose results show that they are applicable to simpler problems with equivalent performance measures. The authors concluded that "if availability constraints come from unexpected breakdowns, fully online algorithms will be needed; but in case of preemptive scheduling, many results of optimality concern the best nearly online algorithms. It is an open question to look for the optimality results from fully on-line algorithms and specific availability patterns, or at least to compute performance bounds." Finally, they stated that a direction is to assume that one operation cannot be interrupted

at any point of time, but only at given instants. Furthermore, they suggested that "authors work on more complicated problems such as sequencing n jobs on m resources in the Flow Job or Open Job environment [11]."

Chaari *et al.* [13] noted that in the real-world several types of hard-to-predict risks must be considered in scheduling problems, and that scheduling under uncertainty allows taking these kinds of risks into account.

They considered that the numerical values (e.g., execution time, machine speed) which are used in the scheduling methods as uncertain, incomplete or imprecise. Consequently, they focused on the dynamic scheduling more than on the deterministic scheduling problems because it is the real scheduling case.

Different types of scheduling approaches under uncertainty environment exist in the literature as well as several dedicated typologies [13]. Most of these typologies are based on the distinction between different scheduling approaches among the following: stochastic, fuzzy [151] proactive and reactive or hybrid which includes two subtypes: predictive-reactive and pro-reactive approaches.

They proposed a global classification schemes technically independent and encompass new kind of scheduling algorithms under uncertainty [13]:

1. Proactive (also known as robust) scheduling approaches: this kind of approach tries to anticipate uncertainty while developing flexibility, in order to produce a schedule, or a family of schedules, that is relatively insensitive to uncertainty [14]. In proactive approaches, five different techniques can be identified: techniques based on robustness measures, redundancy-based techniques, probabilistic methods, contingent scheduling and optimization-based techniques.
2. Reactive scheduling approaches: these approaches are often used in highly perturbed real time environment when off-line scheduling becomes rapidly unfeasible. In this context, decision-making must be very fast and intuitively easy for users to understand [15]. Different methods are used to solve reactive scheduling problems involve distributed (multi-agent systems) or centralized approaches, priority rules and dynamic choice of priority rules. These approaches may exploit priority local criteria when a decision must be made in real-time.
3. Hybrid approaches: these approaches can be subdivided into predictive-reactive and proactive-reactive approaches.

- A. Predictive - reactive approaches: These approaches, used to support risks, have two scheduling phases:
- i) First phase: a deterministic schedule is set up off-line.
  - ii) Second phase: this schedule is used and adapted on-line. The on-line phase requires making scheduling decisions one at a time while the schedule is being built.. These decisions are then adapted in real time to take disturbances into account [16].

Scheduling methods can be constructed by answering two questions: "When to reschedule?" and "How to reschedule?"

- B. Proactive - reactive approaches: In the proactive-reactive approaches, in contrast to predictive-reactive approaches, no rescheduling is done on-line; instead, one among several pre-estimated schedule solutions is chosen. This is what can make it possible to build a set of static schedules such that it is easy, in the event of risk, to pass from one to the other.

"A new mixed technique presented by [17] combines a proactive approach with a reactive approach to deal with scheduling problem under uncertainty. In the proactive phase, the authors built a robust baseline schedule that minimizes the schedules distance defined as the sum of the absolute deviations between the baseline and expected schedules. The robust baseline schedule contains some built-in flexibility in order to minimize the need of complex search procedures for the reactive scheduling approach."

The authors presented an updated and enriched classification scheme for the various approaches for scheduling in an uncertain environment. The more important are as follows [13]. The first is about the notion of predicting uncertain events and measuring their impact on scheduling. The second concerns combining multi-agent approaches with optimization techniques for dynamic scheduling and dynamic control. This prospective has a lack of optimality is due to decisions that lead to the myopic behavior of decision-making entities [18]. A possible solution would be to increase the intelligence of these entities by introducing optimization techniques in the decisional process. The third set of prospective research concerns the integration of new emerging technologies into existing products such as RFID (Radio-Frequency Identification), mechatronics and embedded infotonics. This integration enables the products to participate in the decision-making in a dynamic scheduling context. The fourth set of prospective concerns the new possibilities for designing and dimensioning systems based upon scheduling performances. This takes in account the agility perspective and the issues of evolution and improvement in the production system's (re)design & (re)engineering process.

Kaabi *et al.* [7] focused on results related to the specific real industrial problem of parallel machine scheduling under availability constraints. The authors present parallel machines problems under different constraints with the appropriate solving algorithms along with their complexities under various objective functions. By emphasizing on certain concepts, they pointed out the following:

- Parallel machines scheduling is at hand when machines of similar type and eventually slightly different in characteristics are available in multiple numbers. Jobs can be processed over these machines simultaneously.
- The machines may be subject to accidental breakdowns, periodic preventive maintenance (mainly non-availability), tool changes, workers availability, and availability of the resources used by the machines, and so on.
- There are two classes of problems considered depending on whether the scheduling of preventive maintenance activities is determined before the scheduling of jobs or jointly with the scheduling of jobs. Class 1: in the case that the two activities of the production scheduling and the maintenance “planning” are generally planned and executed separately in real manufacturing systems. Class 2: when maintenance and production services collaborate in order to maximize the system productivity. Maintenance strategies can be broadly classified into Corrective Maintenance (CM) and Preventive Maintenance (PM).
- The job scheduling under maintenance constraints was generally applied to a single machine and multi-machine models. In this paper, the authors deal only with scheduling parallel machines subject to availability constraints.

The authors use multiple tables to represent: Most often used notations in scheduling in Table I, main results of one machine scheduling under availability constraints in Table II and main results on scheduling identical parallel machines under availability constraints in Table III [7]. Subsequently we present two tables summarising the results extracted from this survey concerning the scheduling problems running over uniform and unrelated parallel machines under availability constraints.

TABLE II  
SCHEDULING PROBLEMS RESULTS OVER UNIFORM PARALLEL MACHINES:

Measure to be optimized	Solving Algorithm	Improvement of complexity or competitive ratio	Ref.
$-Q_{1,2} online C_{max}$ (to denote the problem of online scheduling on two uniform parallel machines where one machine is periodically unavailable to minimize the makespan)	Optimal algorithm (as declared by authors )	An online scheduling is investigated on $m$ parallel machines and on two uniform parallel machines where there is one machine periodically unavailable. In the latter case the length of each available period is normalized to 1 while the speed of the other one is $s>0$ . If $s \geq 1$ : speed of the 2nd machine, the proposed algorithm is optimal with a competitive ratio $1+1/s$ . In the case where $0 < s < 1$ , the authors proposed some lower bounds on competitive ratio.	[129]
		If $s=1$ and jobs arrive in decreasing sequence and proved that proposed is optimal with competitive ratio $3/2$ .	[130]

TABLE III  
SCHEDULING PROBLEMS RESULTS OVER UNRELATED PARALLEL MACHINES:

Measure to be optimized	Solving Algorithm	Improvement of complexity or competitive ratio	Ref.
Total machine load on $m$ unrelated parallel machines with maintenance activity (ma)	Two efficient algorithms	Complexity $O(n^{m+3})$	[133]
$Rm  P_{jn}, ma \leq h_{jk}  \sum_{i=1}^m Ci$			
Total machine load			
$Rm  P_{jn} = P_{ij} + ab_{ji}, ma \leq h_{jk}  \sum_{i=1}^m Ci$			
Minimizing the total completion time or the total machine load on $m$ unrelated parallel machines	An algorithm that considers a deterioration of maintenance activities	Complexity $O(n^{m+3})$	[131], [132]
Minimizing the total completion time on $m$ unrelated parallel machines	An algorithm that reconsiders simultaneously a deterioration effects and deteriorating multi-maintenance activities.	Find jointly the optimal maintenance frequencies, the optimal maintenance positions, and the optimal job sequences with a polynomial time algorithm	[134]

The authors concluded that “the maintenance activities can be planned in a flexible or in a non-deterministic ways. In fact, the machines are subject to random breakdowns.” In addition, the assumption of non resumable jobs needs to be taken into account for many real life problems. “Considering more realistic constraints such as online scheduling, resumable jobs, and nondeterministic availability constitute interesting research directions [7].”

The survey paper [19] of Allahverdi provides an extensive review of about 500 papers that have appeared since the mid 2006 to the end of 2014, including static, dynamic, deterministic, and stochastic environments. These survey papers classify scheduling problems based on shop environments as single machine, parallel machine, flow shop, job shop, or open shop. It further classifies the problems as family and non-family as well as sequence-dependent and sequence-independent setup times/costs. In this paper the focus is on the setup times/costs [20] factor ignored by most of the existing scheduling literature. Allahverdi drew up several tables that present the articles references with their criteria or measures to be optimized and their used approaches. These tables are built based on the shop environments as single machine, parallel machine, etc., and according to the problems classification as family and non-family as well as sequence-dependent and sequence-independent setup times/costs. The author summarized the results as follows:

- Heuristics solutions methods have been more used than (the double of) the exact methods.
- The genetic algorithm has been the first one used among the heuristics followed by the Simulated Annealing (SA).
- Among the exact solutions methods, the Mixed Integer Programming (MIP) and the Branch and Bound (B&B) are the most used methods.

He-concluded the need for:

- More research on scheduling problems with explicit consideration of setup times/costs.
- Considering family setup time for the parallel and job shop environments.
- Addressing the sequence-dependent scheduling problems in single machine environments with family setup times.
- Addressing more scheduling problems with multiple criteria.
- Addressing more scheduling problems with uncertain setup times.

In [111] the dynamic job shop scheduling that considers random job arrivals and machine breakdowns was studied. “Considering an event driven policy rescheduling, is triggered in response to dynamic events by variable neighborhood search (VNS). A trained artificial neural network (ANN) updates parameters of VNS at any rescheduling point. Also, a multi-objective performance measure is applied as objective function that consists of makespan and tardiness. The proposed method is compared with some common dispatching rules that have been widely used in the literature for dynamic job shop scheduling problem.”

This paper [119] is characterized by its study of the static scheduling problem of  $m$  identical parallel machines with a common server and sequence dependent setup times. In fact, according to the best knowledge of the authors, it is the first such study in the literature. The authors focused, in their study, on the comparison of the performance of the proposed a Mixed Integer Linear Programming (MILP) model, Simulated annealing (SA) and Genetic Algorithm (GA) based solution approaches with the performance of basic dispatching rules such as: shortest processing time first (SPT) and longest processing time first (LPT) over a set of randomly generated problem instances.

The MILP model is presented with the SPT and LPT dispatching rules for the problem to minimize the makespan. But, according to the authors, the MILP model is not able to solve the large scaled problem instances due to the NP-Hard nature of the problem. For this reason, simulated annealing (SA) and genetic algorithm (GA) based solution approaches are proposed for solving the large scaled problem instances. As a result, based on the computational experiments, the proposed GA is generally the most efficient and effective (followed by the AS approach) in solving this problem. As the problem size increases, the GA approach finds better solutions with smaller standard deviations. Genetic algorithms (GA) are the basis of stochastic optimization algorithms, as they can also be used for machine learning.

Kia *et al.* studied in [6] a dynamic flexible flow line problem with sequence-dependent setup times for minimizing the mean flow time and mean tardiness. By applying genetic programming framework and choosing proper operators, four new composite dispatching rules are proposed to solve this NP-hard problem. To examine scheduling rules performances, a discrete-event simulation model is made considering four new heuristic rules and the six adapted heuristic rules from the literature.

In [117], Rintanen considered that the contingent approach recognizes that different schedules are needed under different contingencies, and computes them either off-line before the execution phase or on-line as information about the contingencies becomes available. This is the most general approach, eliminating the limitations (incompleteness, sub-optimality) of the other approaches at the cost of increased complexity.” He investigated the properties of some classes of contingent scheduling problems. In these problems assignments of

resources to tasks depend on resource availability and other facts that are only known fully during execution. Therefore, the off-line construction of one fixed schedule is insufficient. He demonstrated generally that contingent scheduling is most likely outside the complexity class NP “Their results prove that standard constraint-satisfaction and SAT (Site Acceptance Testing) frameworks are in general not straightforwardly applicable to contingent scheduling.”

Skutella *et al.* address in [120] two main characteristics encountered generally in real-world scheduling problems: the heterogeneous processors and a certain degree of uncertainty about the sizes of jobs. They studied, for the first time according to their best knowledge, a scheduling problem that combines the classical unrelated machine scheduling model with stochastic processing times of jobs.

For the stochastic version of the unrelated parallel scheduling problem and with the objective of weighted sum of completion times  $R |(r_{ij})| \sum w_j C_j$ , they calculated in polynomial time a scheduling policy with a performance guarantee of  $(3 + \Delta)/2 + \varepsilon$  using a novel time-indexed linear programming relaxation. They showed that when jobs also have individual release dates, their bound is  $(2 + \Delta) + \varepsilon$  where,  $\Delta$  is an upper bound on the squared coefficient of variation of the processing times and  $\varepsilon > 0$  is arbitrarily small. They showed that the dependence of the performance guarantees on  $\Delta$  is tight.

On the deterministic side, the current best known approximation algorithms for unrelated parallel machines have respectively performance guarantees  $3/2$  and  $2$  for the problem without and with release dates [121, 122, 124, 125]. Improving these bounds is considered one the most important open problems in scheduling (see Schuurman and Woeginger [126]).

On the Stochastic front, the authors stated that they consider for the first time the stochastic variant of unrelated parallel machine scheduling. In stochastic scheduling, it is asked to compute a non-anticipatory scheduling policy which must make its decisions at an indicated time based on the observed past up to this time as well as the a priori knowledge of the input data of the problem. Here, the processing time of a job  $j$  on machine  $i$  is given by random variable  $P_{ij}$ . The authors assume that the random variables  $P_{ij}$  are stochastically independent across jobs. For any given non-anticipatory scheduling policy, the possible outcome of the objective function  $\sum w_j C_j$  is a random variable. Then, the goal is to minimize its expected value, which by linearity of expectation equals  $\sum w_j \mathbb{E}[C_j]$ .

The authors mentioned that for the first time they completely departed from the linear programming relaxation of Mohring *et al.* [123], and showed how to put a novel, time-indexed linear programming relaxation to work in stochastic machine scheduling. According to the authors, this approach will inspire further research for other stochastic optimization problems in scheduling and related areas. In addition, they showed how to overcome the difficulty that scheduling policies feature a considerably richer structure including complex dependencies between the executions of different jobs which cannot be easily described by time-indexed variables. As a result, they presented the first time-indexed LP relaxation for stochastic scheduling on unrelated parallel machines. Here, they calculated the probability value of a job  $j$  being started on machine  $i$  at time  $t$  which can be represented by the time-indexed variable  $x_{ijt}$ . The situation is complicated in the stochastic context, and it requires a fair amount of information on the exact probability distributions of the random variables. Some other surveys on classical and industrial scheduling problems are summarized in Table IV.

TABLE IV  
PROPERTIES EXTRACTED FROM SOME ADDITIONAL SURVEYS OF CLASSICAL AND INDUSTRIAL SCHEDULING PROBLEMS

Authors / Reference	Scheduling problems Context	Scheduling problems environment	Conclusion and Suggestions
Samia Ourari / [3]	Deterministic to distributed Scheduling approaches based on cooperation	Scheduling approaches under uncertainties: reactive, proactive and proactive-reactive approaches. They differ according to how uncertainty is taken into account, either offline or on-line, single machine & job shop	It is important to address the problem of managing uncertainty in scheduling in order to reduce the gap between the theory (estimated or expected scheduling) and the practical field (scheduling really implemented or adopted).
Janiak et al. / [105]	Offline & online scheduling	Single machine & parallel machine scheduling, Just-in-Time (JIT) scheduling models, PERT/CPM (Program Evaluation Research Technique)/(Critical Path Method) scheduling	A currently noticeable trend in this area is the concept of combining pure due window scheduling problems with other new and trendy phenomena, like e.g., learning or aging effects, deteriorating jobs, maintenance activities, etc. These models need practical trend more than theoretical trend and among their future trends are: the analysis of scheduling problems with due windows (multiple due windows) and preemptive jobs or precedence constraints.

### 3 SCHEDULING PROBLEMS IN CLOUD COMPUTING

Masdari *et al.* [32] addressed scheduling problems in the Cloud environment that, like other environments, still lacks adequate and effective solutions for load balancing and scheduling tasks and workflows. The tasks or

jobs are mapped to the appropriate Virtual Machines (VMs) which are generated virtually from the single physical machine to optimize some given scheduling measures. Various heuristic, metaheuristic and exact algorithms are applied to study the Cloud scheduling problem. In this paper, the authors present an in-depth analysis of the Particle Swarm Optimization (PSO)-based task and workflow scheduling schemes, proposed in the literature, for the cloud environment. Moreover, they provide a classification of the proposed scheduling schemes based on the type of the PSO algorithms illuminating their objectives, properties and limitations. In the Particle Swarm Optimization algorithm, the swarm of particles is randomly generated initially, and each particle position in the search space represents a possible solution and has a fitness value and velocity to determine the speed and direction of its moves. By moving and updating position and velocity, particles get an optimized solution [35, 37]. The authors claim that, following the repeated advances (called iteration), the particle swarm gradually approaches the optimal location [41].

The authors present the papers in literature that propose the PSO algorithms with different schemes:

- In Standard PSO Schemes such as Guo *et al.* [39], Zhang *et al.* [42], Yang *et al.* [38], Huang *et al.* [35], Pandey *et al.* [57], Wu *et al.* [40] and Jianfang *et al.* [36].
- In Multi-Objective<sup>2</sup> PSO Schemes, several contributions were cited in the literature: Netjinda *et al.* [43], Wang *et al.* [139], Ramezani *et al.* [60] and Yassa *et al.* [44].
- In Bi-Objective<sup>3</sup> PSO Schemes; the papers of Beegom *et al.* [48] and Verma *et al.* [50].
- In Hybrid<sup>4</sup> PSO Schemes; the papers of Zhan *et al.* [46], Kaur *et al.* [54], Visalakshi *et al.* [52], Krishnasamy and Gomathi *et al.* [49], Xue *et al.* [53], JieHui *et al.* [55] and Xiaoguang *et al.* [51].
- In Learning PSO Schemes which use learning PSO for scheduling in cloud environment; the papers of Zuo *et al.* [47] and Chen *et al.* [45].
- In Jumping PSO Schemes (which is proposed to optimize the load balancing, the speed-up ratio, and the makespan) the paper of Chitra *et al.* [59].
- In Modified PSO Schemes (it is an improved PSO to overcome the drawbacks of the standard PSO or to increase its performance) the paper of Tarek *et al.* [56], Zhao *et al.* [61], Pragaladan *et al.* [58] and Abdi *et al.* [62].

Properties of these PSO-based scheduling schemes are illustrated in five tables in this survey. Based on these tables, Table V shows in the form of a report a summary of how the schemes studied respond to these properties.

TABLE V  
PROPERTIES REPORT OF SCHEDULING SCHEMES STUDIED IN [32].

PSO Scheduling scheme	References of scheduling schemes studied	The rate of scheduling schemes studied in this survey and which deal with these properties.					
		Objectives			Scheduling type		
		Minimizing cost	Minimizing task execution time	QoS support	Minimizing makespan	Task	Workflow
Standard	[39, 42, 38, 35, 57, 40, 35]	6/7 <sup>f</sup>	-	2/7	4/7	3/7	4/7
Multi Objective	[43, 60, 44, 139]	4/4	2/4	2/4	1/4	3/4	1/4
Bi-Objective	[48, 50]	2/2	1/2	-	1/2	1/2	1/2
Hybrid	[46, 54, 52, 49, 53, 55, 51]	1/7	5/7	1/7	1/7	6/7	1/7
Learning	[47, 45]	-	-	2/2	-	1/2	1/2
Jumping	[59]	-	-	-	1/1	-	1/1
Modified	[56, 61, 58, 62]	3/4	2/4	-	1/4	3/4	1/4

The authors stated that metaheuristic algorithms present better results than deterministic algorithms (with a particular given input, they produce always the same output) in terms of the quality. Likewise, they find approximate solutions faster than traditional exhaustive algorithms in terms of the computation time [138].

Finally, the authors concluded that scheduling is a critical process to map the cloud tasks to the VMs and to reduce their cost and time of execution. Moreover, the future research on scheduling problem should consider, investigate, evaluate and enhance various security related factors in the task and workflow scheduling solutions. As well, the heterogeneous resources functionality, the load balancing on VM (virtual machines) and data center network to reduce their energy consumption. Also, scheduling tasks and workflows on hybrid and federated clouds should be studied more.

<sup>2</sup> This is the process of simultaneously optimizing two or more conflicting objectives subject to a number of constraints.

<sup>3</sup> This is the process which solves optimization models for two objective functions respectively.

<sup>4</sup> To overcome some of the limitations of the PSO, one or more algorithms such as Genetic Algorithm, Ant Colony, etc., are integrated with the PSO algorithm.

<sup>5</sup> e.g., 6 scheduling schemes out of a total of 7 dealt with “Minimizing cost”.

Ramathilagam and Vijayalakshmi [63] noted that there is no typical effective task scheduling algorithm employed in the Cloud environment. Furthermore, the well-known task planners have great difficulty in being implemented in a large-scale distributed environment due to the high communication charge. This requires the building of compatible and applicable job scheduling algorithms and load balancing techniques in this large scale environment. “To balance the load in cloud the resources and workloads must be scheduled in an efficient manner. A variety of scheduling algorithms are used by load balancers to determine which backend server to send a request to [152].” Cloud computing approaches employed the latest technology which is increasing considerably. Job scheduling is one of the process done with the aim to efficiently enhance the functioning of cloud computing atmosphere with achieving maximum profit. In this paper, the authors have investigated and discussed concisely several scheduling algorithms and issues in cloud computing. These algorithms fall into two groups: static and dynamic. Both have their own merits and demerits.

Also, the authors [63] surveyed various types of task scheduling algorithms in Cloud computing. These algorithms are included and compared in Table 1 in Section 5 of their survey. They declared that the heuristic based algorithm, that's belonging to a subset of meta-heuristic approach, is one of the important means to achieve the optimal or near optimal solution of task scheduling in the cloud environment. Many task scheduling techniques that employed in the cloud environment are classified into the following three categories:

1. Traditional techniques which are simple and deterministic, but they get stuck in local optima [67]: First Come First Serve (FCFS), Round Robin (RR) and Shortest Job First (SJF) etc.
2. Heuristic Techniques which are used to find the optimal or near optimal solution by using a sample space of random solutions are: min-min, max-min, enhanced max-min [66] and priority based min-min etc. These techniques give better results as compared to the traditional approaches [69].
3. Meta-heuristic techniques make use of random solution space for tasks scheduling. The principal difference between heuristic and meta-heuristic is the first one is problem specific while the second one is problem independent [70]. Meta-heuristics generally have functional similarities with the aspects of the science of life (biology): (a) Meta-heuristics based on gene transfer: Genetic algorithms and Transgenic Algorithm; (b) Meta-heuristics based on interactions among individual insects: Ant Colony Optimization, Firefly algorithm, Marriage in honey bees Optimization algorithm, Artificial Bee Colony algorithm; and (c) Meta-heuristics based on biological aspects of alive beings: Tabu Search Algorithm, Simulated Annealing algorithm, Particle Swarm Optimization algorithm and Artificial Immune System [153].

Based on this extensive survey, the authors concluded that there remain many problems and issues to enhance as the need of scheduling techniques that covers all requirements accurately. Another issue which is a very vital in scheduling algorithms is the energy efficiency (energy consumption, energy savings, energy sufficiency, etc.). For a multiple workflows, metrics like reliability and availability should also be considered.

Sharma *et al.* describe in [64] the work done in the field of task scheduling algorithms. These algorithms are classified as follows:

1. Efficient Task Scheduling Algorithm: Sindhu *et al.* [65] proposed an enhanced task scheduling algorithm to minimize the completion time of cloudlets. Their approach has two algorithms named as, Longest Cloudlet Fastest Processing Element (LCFP) and Shortest Cloudlet Fastest Processing Element (SCFP).
2. Improved Min-Min Algorithm: Kaur *et al.* [68] proposed an improved min-min algorithm to achieve the maximum resource utilization in distributed environment. This algorithm consists of two phases: the first one is similar to the traditional min-min algorithm, in which minimum completion time of each task is calculated. In the second phase, tasks are rescheduled to make selection of those resources which have been unutilized for a long period of time.
3. Enhanced Max-Min Algorithm: In order to optimize the task scheduling in cloud computing environment, Santosh *et al.* [73] proposed an Enhanced Max-Min Algorithm that consists of two algorithms using respectively the arithmetic and geometric means for calculating average time of job execution instead of maximum completion time. Then the job which has execution completion time just greater than the calculated average time is selected. If jobs are independent of each other, then arithmetic mean gives the best time average. But, on the contrary the geometric mean calculates the best average of time.
4. Selective Algorithm: A selective algorithm has been proposed by Kobra Etminani *et al.* [74] for ensuring the QOS. This algorithm uses the advantages of the two basic scheduling algorithms min-min and max-min and tries to overcome their disadvantages. The selective parameter is the standard deviation of the completion time of unassigned tasks in Meta task.
5. Optimized Task Scheduling Algorithm: for improving scalability in the cloud environment, Shubham Mittal *et al.* [72] proposed an optimized task scheduling algorithm. The authors took into account five algorithms (min-min, max-min, RASA, improved max-min and enhanced max-min).

6. Improved Task Scheduling Algorithm: Abdul Razaque *et al.* [71] proposed an improved task scheduling algorithm to achieve the proper utilization of the network bandwidth in the cloud computing environment. They use a non-linear programming model for assigning a proper number of tasks to each virtual machine.
7. Cloud-based Workflow Scheduling Algorithm: Bhaskar Prasad Rimal *et al.* [75] proposed a Cloud-based Workflow Scheduling Algorithm in order to enhance the workflow in a multi-tenant cloud environment. The authors have defined the Workflow as a new Service layer and the fourth one on the top of the Infrastructure. They used a Directed Acyclic Graph (DAG) to represent the workflow in the systems. The labels on the nodes and the edges represent respectively the costs of computation and of communication. This algorithm uses the ideal time of resources, reduces the makespan, properly utilizes resources and minimizes the cost.
8. Task Scheduling Algorithm based on Quality of Service (QoS): A Task Scheduling – Quality of Service (TS-QOS) algorithm has been proposed by Xiaonian Wu *et al.* [76] for optimizing the service quality in cloud computing. Firstly the algorithm computes the priority of each task on the basis of certain parameters and then it sorts the entire list of tasks according to their priority. The task having minimum completion time is considered as highest priority task and gets the resource first for job completion. Three indexes are taken in account for measuring the performance of this algorithm: i) the makespan, ii) the average waiting time of the longest task (Average Latency), iii) the load balancing index (LBI) used to determine the loading conditions of the system and maximum system loading capacity [154].
9. Task Scheduling Algorithms with Multiple Factors: Nidhi Bansal *et al.* [77] proposed a comparison between traditional scheduling methods, i.e. FCFS, optimization method, QoS-driven, ABC (Activity Based Costing) and priority based algorithms, etc., by using CloudSim as a simulator. The authors showed that the resource utilization and cost factor are the main criteria in any scheduling algorithm to deal at best. Also, they stated that optimization based methods performed better as compared to the traditional methods.

As a result of their study, the authors concluded that there is no such heuristic approach which can fulfill all the required parameters. However, they can perform better when some particular parameter among resource utilization, execution time for each task, and workflow, and so on are considered at a time.

Abbas and Zhang in [83] study the Mobile Edge Computing (MEC), which is an emergent architecture that extends the cloud computing services to the edge of networks leveraging mobile base stations. It can be applied to mobile, wireless and wireline scenarios, using end-users software and hardware platforms located at the network edge.

The mobile networks suffer from low storage and energy capacity, low bandwidth, and high latency [84]. Moreover, exponential growth of the emerging Internet-of-Things (IoT) technology is foreseen to further stumble cellular and wireless networks [85]. The edge computing Fog computing) [86], has begun to be of paramount significance, especially Mobile Edge Computing (MEC) in mobile cellular networks. MEC is equipped with better offloading techniques that characterize the network with low-latency and high-bandwidth.

The basically contribution of this paper is surveying MEC. A few MEC survey reports such as [87] and [88] exist in the literature. It provides a brief overview of different attributes of MEC and identifies the major open research challenges in MEC. In addition, it presents an extensive survey on mobile edge computing focusing on its general overview.

Subsequently, several research efforts were recently carried out in the area of MEC. They are classified according to different domains:

1. Offload computation: This is a way to improve the capacity of mobile devices by transferring the computation to higher resourceful servers that are located at a different location [92]. The improvement mobile devices and networks will still not be able to cope with the increased demand on these devices. As a result, mobile devices will always have to compromise with their limited resources, such as resource-poor hardware, insecure connections, and energy computing tasks [89].

In 2015, many algorithms or prototypes are proposed such as i) Edge Accelerated web Browsing (EAB) prototype proposed by Takahashi *et al.* [91]. It is designed for web applications using a better offloading technique. ii) An algorithm-based design, called Successive Convex Approximation (SCA) proposed by Ardellitti *et al.* [90]. This algorithm optimizes computational offloading on multiple densely deployed radio access points. iii) FemtoCloud system proposed by Habak *et al.* [93] which forms a cloud of orchestrated co-located mobile devices that are self-configurable into a correlative mobile cloud system

In 2016, other algorithms or prototypes are proposed: a) the efficient computation offloading model designed by Chen *et al.* [94] using a game theoretic approach in a distributed manner. Game theory is a persuasive tool that helps simultaneously connected users to make the correct decision when connecting a wireless channel based on the strategic interactions. b) the contract-based computation

resource allocation scheme proposed by Zhang *et al.* [95]. It improves the utility of vehicular terminals which intelligently utilize services offered by MEC service providers under low computational conditions.

2. Low Latency: MEC is equipped with better offloading techniques that characterize the network with low-latency and high-bandwidth. So, it is one of the promising edge technologies that can improve user experience by providing high bandwidth and low latency.

In 2015, Nunna *et al.* [97] proposed a real-time context aware collaboration system by combining MEC with 5G networks. This integration of MEC and 5G helps to empower real time collaboration systems utilizing context-aware application platforms. These systems require context information combined with geographical information and low latency communication.

In 2016, two schemes are proposed: i) REPLISOM designed by Abdelwahab *et al.* [96] is an edge cloud architecture and (Long Term Evolution) LTE enhanced memory replication protocol to avoid latency issues. LTE bottleneck occurs when allocating memory to a large number of IoT devices in the backend cloud servers. ii) Kumar *et al.* [99] proposed a vehicular delay tolerant network-based smart grid data management scheme. The authors investigated the use of Vehicular Delay-Tolerant Networks (VDTNs) to transmit data to multiple smart grid devices exploring the MEC environment.

3. Storage: To overcome their device storage limitation, end-users may utilize MEC resources.

In 2016, Jararweh *et al.* [98] proposed a framework connects software defined system components to MEC to further extend MCC (Mobile Cloud Computing) capabilities, Software Defined system for Mobile Edge Computing (SDMEC). The components jointly work cohesively to enhance MCC into the MEC services.

4. Energy Efficiency: The MEC architecture is created to reduce energy consumption of user devices by migrating compute intensive tasks to the network edge.

Many schemes are developed in this field:

In 2014, an opportunistic peer-to-peer mobile cloud computing framework was proposed by Wei Gao [103]. The probabilistic framework is composed of peer mobile devices connected via their short-range radios. Based on their available capacity, these mobile devices are able to share energy and computational resources. The author proposed the probabilistic method to estimate the opportunistic transmission status of the network ensuring that the resulting computation is timely delivered to its initiator.

In 2015: i) an architecture that integrates MEC to voice over LTE called ME-VoLTE was proposed by Beck *et al.* [101]. The encoding of video calls is offloaded to the MEC server located at the base station (eNodeB). The offloading of video encoding through external services helps escalating battery lifetime of the user equipment. Encoding is high computational-intensive and hence is very power consuming. ii) El-Barbary *et al.* [100] proposed DroidCloudlet; an architecture to enhance mobile battery lifetime by migrating data-intensive and compute-intensive tasks to rich-media. Based on commodity mobile devices DroidCloudlet is legitimized with resource-rich mobile devices that take the load of resource-constraint mobile devices. iii) in 2016 Jalali *et al.* [102] proposed a flow-based and time based energy consumption model. They conducted number of experiments for efficient energy consumption using centralized nano Data Centers (nDCs) in a Cloud computing environment. But, the authors claim that nDCs energy consumption is not yet been investigated.

The authors [83] concluded that as a recent technology platform, little research has been specifically done in MEC. In fact, there are some open issues in MEC that need to be addressed. Many researchers interested by MEC have studied some problems of these issues that belong in several areas including: Resource Optimization, Transparent Application Migration, Web Interface, Security, Pricing, Network Openness, Multi-services and Operations, Robustness and Resilience.

We summarise, in Table VI, significant properties extracted from some other surveys of scheduling problems in Cloud environment: context, environment, conclusion and suggestions

TABLE VI  
ADDITIONAL SURVEYS OF SCHEDULING PROBLEMS IN CLOUD ENVIRONMENT

Authors / TReference	Scheduling problems Context	Scheduling problems environment	Conclusion and Suggestions
S. Yi <i>et al.</i> / [106]	Static, Dynamic, Real Time and Heuristic Scheduling	The main issues: Fog networking (is heterogeneous): Internet of Things, software-defined networking, network function virtualization (NFV) to create flexible and easy maintaining network environment. Quality of Service (QoS): connectivity, reliability, capacity, and delay.	Fog computing will evolve with the rapid development in underlying IoT, edge devices, radio access techniques, SDN, NFV, VM and Mobile cloud. We think fog computing is promising but currently need joint efforts from underlying techniques to converge to "fog computing".

		Interfacing and programming model. Computation Offloading. Accounting, billing and monitoring. Provisioning and resource management: Application-aware provisioning, Resource discovery and sharing. Security and Privacy	
Deshmane <i>et al.</i> / [107].	Static, Dynamic, Real Time and Heuristic Scheduling	Enhanced Max-min, Improved Genetic, Scalable Heterogeneous Earliest-Finish-Time (SHEFT) Improved Cost-Based, Resource-Aware-Scheduling, Innovative transaction intensive cost-constraint scheduling, Algorithms and Multiple QoS Constrained Scheduling Strategy of Multi-Workflows (MQMW)	There is a need to implement a new scheduling algorithm to minimize the execution time and improve availability and reliability in a cloud computing environment. The improvement can also be done with building algorithms that take user preferences while scheduling. Also, one more aspect can help improving the design of algorithm, which can include new factors such as inter-node bandwidth etc., that have not been considered for resources matching.
Ahmed <i>et al.</i> / [88]	Static, Dynamic, Real Time and Scheduling	The contributions of this article are as follows: (a) survey of the state-of-the-art research. (b) Preview taxonomy based on various parameters such as characteristics, actors, access technologies, applications, objectives, computational platforms, and key enablers. (c) Identification of various open challenges related to the Mobile Edge Computing that impede or prevent the successful deployment.	The open research challenges in Mobile Edge Computing: Standard Protocol, Simulation Platform, Mobility Management, Heterogeneity, Pricing Model, Scalability and Security
Singh <i>et al.</i> / [109]	Dynamic / static scheduling and allocation of resources	Resource scheduling algorithms (RSA) and dynamic RSAs, Bargaining Based RSA, Compromised Cost and Time based RSA, Cost Based RSA, Dynamic and Adaptive Based RSA, Energy Based RSA, Hybrid Based RSA, Nature Inspired and Bio-Inspired Based RSA, Optimization Based RSA, Profit Based RSA, Priority Based RSA, SLA and QoS Based RSA, Time Based RSA and VM Based RSA  Resource scheduling aspects and resource distribution policies	Recent research has shown that resource scheduling algorithms using resource provisioning mechanisms and applying the effective resource provisioning technique.  On the basis of existing research, it is necessary to fully understand QoS requirements for workload for better allocation of resources rather than to detect workload and resources. It is necessary to find the progress in the search on the cloud itself before finding the advanced search in the scheduling of resources.
Manpreet Kaur / [110]	Static & Dynamic scheduling	Min-Min, Max-Min, RASA, Shortest Job First, heuristic, Dominant Resource Priority and multi-objective task Scheduling Algorithms	Future work would be to continue the multi-objective scheduling improvement. The authors have done the non-dominated sorting of virtual machines (VMs) according to MIPS. In future they aim to take other parameters of VMs also to sort them for better performance.

#### 4 SCHEDULING PROBLEMS IN MOBILE NETWORKS

Many surveys have been published recently with regard to scheduling problems in mobile networks. In this section, we address these surveys briefly.

Mahidhar *et al.* [21] addressed the problem of the wireless sensor network (WSN) by providing dynamic scheme. WSN is a highly distributed network of small and light nodes. This problem is demonstrated by the limited battery life of the nodes. Sensor nodes spend their energy in transmitting and receiving the data, as well as, in the relaying of the packets. This implies designing the routing algorithm that maximizes the lifetime of the network. Packet scheduling is important in WSN to maintain fairness based on data priority and to reduce end-to-end delay. The authors proposed the Dynamic Multilevel Priority (DMP) Packet Scheduling Scheme with the Bit Rate classification; the data is divided into three categories as high, moderate and low bit rate. They also proposed the threshold value check mechanism to prevent deadlock situations. To provide security they implement the RC6 security algorithm.

Another important implication in real-time WSN data transmission is the packet scheduling at node sensors that ensures the delivery of different packages according to their priority and fairness without any delay. This saves battery power by reducing sensors working time.

The authors also presented various existing real time scheduling schemes which are as follows [21]:

1. Dynamic Conflict Free Transmission Scheduling (DCQS): is a query based novel scheduling technique, designed to support in network data aggregation and in response to the workload changes it can dynamically adapt to the transmission [23, 26].

2. Nearest Job Next (NIN): It consists of the mobile element (ME); server and client. The client is the one which request the service and it is a simple and intuitive discipline which is adopted by the ME to select the next request to be served or the next client [22].
3. Traffic Pattern Oblivious Scheduling (TPO): to handle efficiently a wide variety of the traffic pattern by using a single TDMA (Time-Division Multiple Access) schedule [27].
4. Dynamic Multilevel Priority Packet Scheduling (DMPPS): it consists of three levels of the priority queues, the data is placed in the priority queue based on the priority, the last level of the virtual hierarchy does not have the priority queue and the levels are formed based on the hop distance from the base station [25].
5. First Come First Serve (FCFS): it is the simplest packet scheduling algorithm in which packets are processed as they come [24].

By comparing these schemes, the authors concluded that DMPPS is the better one. They presented a literature review table for five articles [25,29,31,28,30]. This table shows the objective, key issues, and advantages which are summarized by: reducing average energy consumption, balancing the nodes energy consumption, minimizing the delay at nodes and increasing network life. They also illustrated the Adaptive Staggered Sleep Protocol ASLEEP protocol which is efficient for the power management in wireless sensor network. This protocol adjusts dynamically the node sleep schedulers to match the network demand. The node adjusts its active period dynamically [21, 137].

The scheduling scheme proposed in [21] has three levels of the priority queues; the last level of the virtual hierarchy does not have the priority queue. The data packet classification is done as i) real time data given as priority 1, ii) non real time remote data, received from the lower level nodes, given as priority 2, iii) non real time local data, sensed from the node itself, given as priority 3. The TDMA scheme is used to process the data packet sensed by the node which are at the different levels. The conclusion drawn from this paper is that one of the advantages of their DMP scheme is its dynamicity to the changing requirement of the Wireless Sensor Network application. The proposed threshold value, to check mechanism at the time of the priority level when the data arrives at the high priority queue, helps to reduce the deadlock situation.

In his survey [78], Nimbalkar introduces Opportunistic Scheduling for effective load balancing in multipath traffic network. This is a technique that aims to maximize throughput and packet delivery ratio by exploiting short-term variation in path condition.

Opportunistic Scheduling works to achieve two objectives simultaneously: selecting the user with the best channel conditions and satisfying the fairness constraints over long-term scales. Thus, many algorithms are developed, as in [79, 80, 81, 82] which exploit high-quality channels to realise the fairness use of the multiple channels available.

Opportunistic Scheduler takes into account some criteria such as maximum delay, minimum throughput, maximum response time, maximum latency in the way to impose fairness on the channels. So, certain characteristics should be contained in a good scheduling algorithm such as maximum resource utilization, maximum throughput, minimum turnaround time, minimum waiting time, and minimum response time.

The Opportunistic Scheduling techniques has been studied and classified in five categories. The following table (Table VII) summarizes their characteristics.

TABLE VII  
CHARACTERISTICS OF OPPORTUNISTIC SCHEDULING FOR TRAFFIC NETWORKS WITH MULTIPLE CONSTRAINTS [78]

Category	Approach	Optimized measures	Channels / scheduling Categories	Performance	Resolution settings	Ref.
Fairness	Proportional-fair sharing approach	Total throughput	Two competing channels	Maximize total throughput	The channel rate history	[140]
	Multichannel Fair Scheduler	Optimal throughput	Multiple wireless channels / Deterministic and Probabilistic	Maximize throughput	This model uses adaptive control framework to develop opportunistic fair wireless schedulers.	[141]
	Indexed to optimal solution of throughput	QoS and throughput	Multiple user QoS	Optimal solution of throughput		[79]
	Heuristic opportunistic scheduling policy	QoS, throughput and short-term fairness	Multiple interface system	Throughput performance of the heuristic policy is comparable to that of the long- term optimal policy	Heuristic opportunistic scheduling policy	

Delay	This model uses Lyapunov optimization framework for stochastic network optimization	Stronger delay & efficient throughput utility	Stochastic	Opportunistic scheduling that guarantees a bounded worst-case delay	Network which has time-varying channel condition	[80]
	Using radio resource allocation in OFDMA	Delay sensitive traffic, system capacity and throughput fairness		This model provides fairness with respect to the realizable throughput per user, packet dropping ratios and packet delay distributions	OFDMA (Orthogonal frequency-division multiple access) based network	[81]
	An opportunistic scheduling based on multi-user diversity effect	Total throughput		Total system throughput is maximized	This scheduling mechanism can result in higher spectrum utilization	[82]
Quality of Service	A time slotted system where time is the resource to be shared among all users	Maximize the system performance stochastically	Multiple channels / Stochastic	Users scheduling, to transmit slots at each time, that optimize the network performance.		[108]
	Model uses a time-slotted system where time is the resource to be shared	Throughput value	Multiple channels / Stochastic	The performance of a user's channel condition by enlarging the stochastic process value.		[142]
	A model whereby the scheduling mechanism is based on preventing from transmitting under adverse conditions.	Maximize user utility measure, e.g., communication rate for efficient utilization of the available communication resources.	Shared wireless channels	Communication rate	Resources of a wireless channel network	[143]
	Reinforcement learning framework to design distributed adaptive opportunistic routing problem (d-AdaptOR)	Minimizing the average per packet cost for routing a packet from source to destination	Wireless multi-hop network	Packet routing from source to destination	Distributed adaptive opportunistic routing problem (d-AdaptOR)	[144]
Throughput	Low complexity adaptive scheduling algorithms	An identical throughput guarantee	Time-slotted networks	An approximate throughput guarantee	This model develops an expression for the approximate throughput guarantee violation probability for users in time-slotted networks	[145]
	A simple algorithm for networks with short-lived flows	Throughput optimal	Wireless networks with short-lived flows	Performance of the channel flows transmission	Wireless channel network	[146]
	Opportunistic Multipath Scheduling (OMS)		Multipath routing uses multiple alternative paths in the network.	OMS minimize the delay and improves overall throughput	Multiple network paths	[147]
.	A new model of opportunistic scheduling mechanism	Maximizing the system overall throughput	Wireless network with hybrid links	To avoid starvation of the link having a much lower transmission rate	Wireless channel network	[148]
	A model of Distributed Opportunistic Scheduling (DOS) under average delay	Maximize the overall throughput or the throughput of every link according to its own individual	Ad-hoc network of wireless channels	Optimize the throughput performance in Ad-hoc network	Homogeneous/heterogeneous scenarios with saturated/non-saturated stations	[149]

	constraint	delay constraint				
	Distributed Opportunistic Scheduling (DOS)	Maximize the throughput in a network of channels with a certain access probability	Wireless network	Improve the throughput performance of the wireless network	Homogeneous/heterogeneous scenarios with saturated/non-saturated stations	[150]

In [112], the authors write "we present a centralized integrated approach for: 1) enhancing the performance of an IEEE 802.11 infrastructure wireless local area network (WLAN), and 2) managing the access link that connects the WLAN to the Internet. Our approach, which is implemented on a standard Linux platform, and which we call ADvanced Wi-fi Internet Service EnhanceR (ADWISER), is an extension of our previous system WLAN Manager (WM). ADWISER addresses several infrastructure WLAN performance anomalies such as mixed-rate inefficiency, unfair medium sharing between simultaneous TCP uploads and downloads, and inefficient utilization of the Internet access bandwidth when Internet transfers compete with LAN-WLAN transfers, etc. The approach is via centralized queueing and scheduling, using a novel, configurable, cascaded packet queueing and scheduling architecture, with an adaptive service rate."

In the objective of managing inter-cell interference with a centralized controller, Ramos-Cantor *et al.* addressed in [116] the problem of coordinating scheduling decisions among multiple base stations in an -LTE-Advanced downlink network. In order to solve the coordinated scheduling problem an integer non-linear program is formulated. It only makes use of the specific measurement reports defined in the 3GPP standard. Unlike most existing approaches, it does not rely on exact channel state information. The authors proposed an equivalent integer linear programming reformulation of the coordinated scheduling problem, which can be solved efficiently by commercial solvers. The performance of the proposed coordinated scheduling approaches is analyzed by extensive simulations of medium to large size networks. The available analytical results show fundamental limits in cooperation due to interference outside the cluster.

A centralized scheduling approach was proposed in [114] to manipulate centralized coordination among heterogeneous agents. In this study, the center agent acts as an information collector, processor and resource scheduler. The main contribution of this center is enacts centralized scheduling to run well. A clustering analysis based on artificial immune algorithm is applied to process information, moreover a series of schemes are suggested to ensure smooth scheduling.

In [113], it is indicated that the data-scale computing for analytical workloads is becoming increasingly popular. Due to the high operational costs heterogeneous applications are forced to share cluster resources to achieve economies of scale. Existing approaches of scheduling large and diverse workloads are tackled in two alternative ways: (1) the solutions offer strict, secure enforcement of scheduling invariants (fairness, capacity) for heterogeneous applications; and (2) the distributed solutions offer scalable, efficient scheduling for homogeneous applications. The authors proposed Mercury, a hybrid resource management framework that supports the full spectrum of scheduling, from centralized to distributed. Mercury exposes a programmatic interface that allows applications to trade-off between scheduling overhead and execution guarantees. The authors stated that their framework harnesses this flexibility by opportunistically utilizing resources to improve task throughput. Also, experimental results on production-derived workloads show gains of over 35% in task throughput. These benefits can be translated by appropriate application and framework policies into job throughput or job latency improvements. They have implemented and contributed Mercury as an extension of Apache Hadoop / YARN.

Fu *et al.* introduced in [115] En-Omega. This is a novel hierarchical hybrid design of schedulers to address the serious job starvation problem that triggers especially in heavily loaded clusters. In En-Omega the fully distributed schedulers can be enhanced with a central scheduler. This can provide a global fairness to the jobs from different schedulers and simultaneously reduce the average latency of all the jobs sharply. To reduce the overhead, in En-Omega design, the central scheduler will be activated only when the cluster is heavily loaded. Furthermore, the cache used for central queuing and the scoring policy used in central scheduling are all load-aware. En-Omega was evaluated based on Google trace and experimental results show that, compared to the baseline design, this method can reduce the average latency of starving jobs up to 90% with reasonable overhead.

In [118], the authors pointed out that the traditional distributed wireless video scheduling is based on perfect control channels where instantaneous control information from the neighbors is available. They mentioned the difficulty to obtain this information in practice, especially for dynamic wireless networks. They found that the two approaches - distortion-minimum scheduling aiming to meet the long term video quality demands and the

other that focusing on a minimum delay - can't be applied directly. Then they went to investigate the distributed wireless video scheduling with delayed control information (DCI). First, they translate this scheduling problem into a stochastic optimization rather than a convex optimization problem in the way to exploit in a tractable framework. Next, they consider two classes of DCI distributions to study the relationship between the DCI and scheduling performance, and provide a general performance property bound for any distributed scheduling. These classes are: i) the class with finite mean and variance, and ii) a general class that does not employ any parametric representation. Thereafter, it will be created a class of distributed scheduling scheme to achieve the performance bound by making use of the correlation among the time-scale control information. The main contributions are presented at the theoretical and technical level.

1. Theoretical level: an appropriate Lyapunov function based on observed DCI is presented to establish the scheduling performance in terms of DCI. This leads to the positive Harris recurrence property of the network Markov process. This represents the most challenging part of this work since it needs to prove an effective time scale separation between the network state dynamics and scheduling decision dynamics. To make this possible, they design an increase function of queue-size to capture the correlation of the DCI.
2. Technical level: a distributed video scheduling scheme in terms of DCI is proposed. This scheme only utilizes local queue-length information to make scheduling decisions. Also, it only requires each node to perform a few logical operations at each scheduling decision.

The authors concluded that they provided a general performance property bound for any distributed scheduling. Importantly, they designed a class of distributed online scheduling scheme to achieve the optimal performance bound by making use of the correlation among the time-scale control information.

Some other surveys of scheduling problems in mobile networks, whose important characteristics are summarized in Table VIII, are: Context, Environment, Conclusion and Suggestions.

TABLE VIII  
PROPERTIES EXTRACTED FROM SOME ADDITIONAL SURVEYS OF MOBILE NETWORKS SCHEDULING PROBLEMS

Authors / Reference	Area	Context	Scheduling problems environment	Conclusion and Suggestions
Akashdeep <i>et al.</i> / [104]	Mobile networks	Point-to-Multipoint (PMP) and mesh mode for wireless broadband access in networks.	Scheduling techniques for IEEE 802.16 networks in PMP Mode and their approaches which may be divided into sub categories such as Traditional, Hierarchical, Cross Layer Approaches, Dynamic Schedulers and Soft Computing based.	<p>There are some of the areas still not quite explored namely the application of soft computing/optimization techniques like Genetic Algorithm, neural networks, fuzzy logic etc.</p> <p>Using these approaches together with information from higher layers can act as a major contributor in the field of scheduling.</p>

## 5 DISCUSSION

Through recent surveys presented in the present paper and others more ancient, we conclude that the scheduling problem has been developed by balancing tasks on one machine, on a few machines and then over very large number of machines. This is developed with time, over decades, from heavy manufacturing to various areas of light industries passing recently in networks of mobile devices going to Cloud, Fog and Edge computing.

There are many different algorithms designed by researchers or adopted from several domains, especially from mathematics, and then developed to solve scheduling problems. In this paper, we are investigating their advantages in terms of efficiency and speed of achieving the desired optimal scheduling policy.

Based on what was published, we find that the criterion to meet is to accomplish the desired job efficiently in the shortest possible time, taking into account the capacity of equipment used in terms of energy and treatment. It is obvious that researchers prefer an exact optimal policy as a resolution approach to solving scheduling problems. But, if this policy has a great complexity and requires more effort, energy and time, researchers aim to find another policy close to it simpler and more economy. This is the case of small mobile devices in networks that support the latest technologies, for example Cloud and Fog Computing. Especially, for the communication systems of ubiquitous resources that have several challenges still cannot be accommodated such as: high energy consumption, the continuously increasing demand for high data rates and the mobility required by new wireless applications. So, according to the fact that these devices suffered of poor energy and storage or computing capacity, the trend is the use of heuristic and meta-heuristic methods.

We mention the success achieved through the use of Cloud computing in solving the problems of communication as well as the storage and transfer of data. Then, the use of Fog and Edge computing to solve the problems experienced by Cloud computing. But, we find that these means become ineffective in the case of

heavy and excessive communications or the case of natural or terrorist disasters that may destroy their service centers. Therefore, we believe that the solution in these cases is the network of ubiquitous resources and mobile devices that can be established anytime and anywhere. This can also be used in normal situations because of their very low cost and their permanent presence with us.

## 6 CONCLUSION

Scheduling and load balancing problems are the most important issues for sharing tasks between machines or devices. The most recent systems, such as Internet objects, cloud computing, fog and edge, whose main issue is sharing tasks, are rapidly evolving because of technological advances and excessive demand for their use.

In this paper, we presented a collection of recent surveys and articles that are interested in scheduling problems. We presented different scheduling issues, and investigated their advantages and disadvantages based on many criteria, such as context, environment, optimizing function, used algorithms, suggestions and proposed improvement. This may give rise to a document which summarizes most recent researches in this field in order to present a comprehensive idea of scheduling problems.

The main objective of this effort is to facilitate the work of researchers and readers who investigate this model as well as those working in this field. Also, to attract attention to enhance and thus take advantage of newly emerging areas that are in dire need of scheduling such as Cloud, Fog and Edge computing, especially the network of ubiquitous resources and mobile devices.

As a result, we have concluded that efforts in this area focus on optimizing the desired measures and achieving jobs in the shortest possible time. But, this seems to be unattainable by using exact and optimal policies of scheduling and load balancing tasks on the available devices. This is due to the algorithms complexity that gives these exact solutions. Therefore another way is adopted, in the literature, by means of heuristic and approximate algorithms which search optimal policies close to the exact ones, which decreases considerably the execution time. We recommend and urge to focus the future works to develop this type of algorithms.

Finally, we aim to attract researchers to enhance scheduling issues in the individual mobile networks segment of collaborative mobile computing. The main factor behind this trend is to create a practical alternative to the applications of information, communications, task scheduling and many more, for example in cases of disasters and terrorism when it becomes impossible to move from place to another and use the centers dedicated to these applications. Another reason, to be realized, is the facilities and benefits that these issues can provide to users in terms of implementing and executing their applications, and thus managing their various jobs, through mobile networks at any time and wherever they are. This is for nowadays, but for years to come, we claim that they will be the de facto devices due to the anticipated future technological improvements.

## REFERENCES

- [1] J. Leung, J. Anderson, "Handbook of Scheduling: Algorithms, Models, and Performance Analysis," CRC/Chapman-Hall, 2004.
- [2] T. Vredenveld, "Stochastic online scheduling", August 2012, Volume 27, Issue 3, pp 181–187.
- [3] S. Ourari, "De l'ordonnancement déterministe à l'ordonnancement distribué sous incertitudes 2011," Thèse de doctorat, Université Paul Sabatier de Toulouse, France, 2011.
- [4] K. R. Baker, "Introduction to sequencing and scheduling," Wiley, New York, 1974.
- [5] M. Pinedo, "Scheduling: Theory, algorithms and systems," 2<sup>nd</sup> edition, Prentice Hall, 2008.
- [6] H. Kia, S. H. Ghodspour, H. Davoudpour, "New scheduling rules for a dynamic flexible flow line problem with sequence-dependent setup times," Journal of Industrial Engineering International, 2017.
- [7] J. Kaabi, Y. Harrath, "A survey of parallel machine scheduling under availability constraints", International Journal of Computer and Information Technology, 3(2):238–245, 2014.
- [8] Y. Ma, C. Chu, C. Zuo, "A survey of scheduling with deterministic machine availability constraints", Computers & Industrial Engineering, Vol 58, p. 199-211, 2010.
- [9] Z. Tan, Y. He, "Optimal online algorithm for scheduling on two identical machines with machine availability constraints," Information Processing Letters, 83, 323–329. 2002.
- [10] P. Senthilkumar, S. Narayanan, "Literature Review of Single Machine Scheduling Problem with Uniform Parallel Machines," Intelligent Information Management, 2010, 2, 457-474.
- [11] H. R. D. Saidy, M. T. Taghavi-Fard, "Study of Scheduling Problems with machine Availability Constraint," Journal of Industrial and Systems Engineering, Vol. 1, No. 4, p. 360-383, 2008.
- [12] J. Blazewicz, J. Breit, P. Formanowicz, W. Kubiak, G. Schmidt, "Heuristic algorithms for the two-machine flowshop problem with limited machine availability," Omega Journal 29; 599-608, 2001.
- [13] T. Chaari, S. Chaabane, N. Aissani, D. Trentsaux, "Scheduling under uncertainty: Survey and research directions", In Proceedings of Advanced Logistics and Transport (ICALT), pages 229–234. IEEE, 2014.
- [14] X. Lin, S. L. Janak, C. A. Floudas, "A new robust optimization approach for scheduling under uncertainty: I. Bounded uncertainty," Computers & Chemical Engineering, FOCAP 2003 Special issue, 28 (6-7), 1069-1085, 2004.
- [15] E. Szelke, R. Kerr, "Knowledge-based reactive scheduling," Production Planning & Control, 5(2), 124, 1994.
- [16] P. Cowling, D. Ouelhadj, S. Petrovic, "Dynamic scheduling of steel casting and milling using multi-agents", Production Planning & Control, 15(2), 178-188, 2004.
- [17] C.W. Wu, K.N. Brown, J.C. Beck, "Scheduling with uncertain durations: Modeling Beta-robust scheduling with constraints," Computers and Operations Research, 36, 2348-2356, 2009.

- [18] D. Trentesaux, "Distributed control of production systems," *Engineering Applications of Artificial Intelligence* 22(7), 971-978, 2009.
- [19] A. Allahverdi, "The third comprehensive survey on scheduling problems with setup times/costs", *Euroepan Journal of Operational Research*, 246, 345–378, 2015.
- [20] A. Allahverdi, H. M. Soroud, "The significance of reducing setup times/setup costs", *European Journal of Operational Research*, 187, 978–984, 2008.
- [21] R. Mahidhar and A. Raut, "A Survey on Scheduling Schemes with Security in Wireless Sensor Networks", *Procedia Computer Science* Journal, 78, 756-762, 2016.
- [22] L. He, Z. Yang. "Evaluating Service Disciplines for On-Demand Mobile Data Collection in Sensor Networks," *IEEE Transactions on mobile computing*, vol. 13, no. 4; April 2014.
- [23] R. Gomathi, N.Mahendran, "An Efficient Data Packet Scheduling Scheme in Wireless Sensor Networks", *IEEE 2nd international conference on electronics and communication system (ICECS)*, 2015.
- [24] G. D. Celik, E. H. Modiano, "Controlled mobility in stochastic and dynamic wireless networks," *IEEE CDC'10*; May 2012, 251–277.
- [25] N. Nasser, L. Karim, T. Taleb, "Dynamic Multilevel Priority Packet Scheduling Scheme for Wireless Sensor Network", *IEEE Transactions on wireless communications*. Vol.12, no.4; april 2013.
- [26] J. A. Stankovic, C.-G. Roman, "Dynamic Conflict-Free Transmission Scheduling for Sensor Network Queries", *IEEE Transactions on mobile computing*, 10(5), May 2011.
- [27] A. Somasundara, A. A. Ramamoorthy, "Mobile Element Scheduling for Efficient data collection in wireless sensor network with Dynamic Deadline," *Department of Electrical Engineering, IEEE International Real-Time Systems Symposium*, 1052-8725, 2004.
- [28] G.Bergmann, M.Molnar, L.Gonczy, B. Cousin, "Optimal Period Length for the CQS Sensor Network Scheduling Algorithm," in *Proc. 2010 International Conf. Netw. Services*, 192-199.
- [29] B. Nazir and H. Hasbullah, "Dynamic Sleep Scheduling for Minimizing Delay in Wireless Sensor Network", in *Proc.2011 Sadi International Electron, Communications Photon*, Conf.,1-5.
- [30] A.R.Swain, R.C.Hansdah, and V.K.Chouhan, "An Energy Aware Routing Protocol with Sleep Scheduling for Wireless Sensor Networks," in *Proc. 2010 IEEE International Conf. Adv. Inf. Netw. Appl*, 933-940.
- [31] F.Liu, C. Tsui, and Y. J. Zhang, "Joint Routing and Sleep Scheduling for Lifetime Maximization of Wireless Sensor Networks", *IEEE Trans. Wireless Communication*, 9(7), 2258-2267, 2010.
- [32] M. Masdari, F. Salehi, M. Jalali, and M. Bidaki, "A survey of PSO based scheduling algorithms in cloud computing", *J. Netw. Syst. Manage.*, 25(1), 122–158, 2016.
- [33] J. Kennedy, R. C. Eberhart, "Particle Swam Optimization", *Roc. IEEE International Conference on Neural Networks (Path, Australia)*, IEEE Service Center, Piscataway, NJ, N: 1942-1948, 1995.
- [34] P. Shelokar, P. Siarry, V. K. Jayaraman, B. D. Kulkarni, "Particle swarm and ant colony algorithms hybridized for improved continuous optimization", *Appl. Math. Comput.* 188, 129–142, 2007.
- [35] J. Huang, K. Wu, L. K. Leong, S. Ma, M. Moh, "A tunable workflow scheduling algorithm based on particle swarm optimization for cloud computing", *Criterion*, 12-14, 2013.
- [36] C. Jianfang, C. Junjie, Z. Qingshan, "An optimized scheduling algorithm on a cloud workflow using a discrete particle swarm," *Cybern. Inf. Technol.* 14, 25–39, 2014.
- [37] R.C. Eberhart, Y. Shi, "Particle swarm optimization: developments, applications and resources," In: *Evolutionary Computation, 2001. Proceedings of the 2001 Congress on ????*, 81–86, 2001.
- [38] X. Qin, Z. Yang, W. Li, Y. Yang, "Optimized task scheduling and resource allocation in cloud computing using PSO based fitness function," *Inf. Technol. J.*, 12, 7090–7095, 2013.
- [39] L. Guo, S. Zhao, S. Shen, C. Jiang, "Task scheduling optimization in cloud computing based on heuristic algorithm", *J. Netw.* 7, 547–553, 2012.
- [40] Z. Wu, Z. Ni, L. Gu, X. Liu, "A revised discrete particle swarm optimization for cloud workflow scheduling," In: *Computational Intelligence and Security (CIS), International Conference on ????*, 184–188, 2010.
- [41] X. Wen, M. Huang, J. Shi, "Study on Resources Scheduling Based on ACO Algorithm and PSO Algorithm in Cloud Computing," In: *Distributed Computing and Applications to Business, Engineering & Science (DCABES)*, 2012 11th International Symposium on ????, 219–222, 2012.
- [42] G. Zhang, X. Zuo, "Deadline constrained task scheduling based on standard-PSO in a hybrid cloud", In: *Advances in Swarm Intelligence*, 200–209. Springer, 2013.
- [43] N. Netjinda, T. Achalakul, B. Sirinaovakul, "Cloud provisioning for workflow application with deadline using discrete PSO", *ECTI Trans. Comput. Inf. Technol.* 7, 43–51, 2013.
- [44] S. Yassa, R. Chelouah, H. Kadima, B. Granado, "Multi-objective approach for energy-aware workflow scheduling in cloud computing environments," *Sci. World J.* 2013, 350934 (2013). doi:10.1155/2013/350934.
- [45] W.-N. Chen, J. Zhang, "A set-based discrete PSO for cloud workflow scheduling with user-defined QoS constraints," In: *Systems, Man, and Cybernetics (SMC), 2012 IEEE International Conference on*, 2012, pp. 773–778.
- [46] S. Zhan, H. Huo, "Improved PSO-based task scheduling algorithm in cloud computing," *J. Inf. Comput. Sci.* 9, 3821–3829 (2012).
- [47] X. Zuo, G. Zhang, W. Tan, "Self-adaptive learning PSO-based deadline constrained task scheduling for hybrid IaaS cloud," *IEEE Trans. Autom. Sci. Eng.* 11, 564–573 (2014).
- [48] A.A. Beegom,M. Rajasree, "A particle swarm optimization based pareto optimal task scheduling in cloud computing," In: *Advances in Swarm Intelligence*, pp. 79–86. Springer (2014).
- [49] K. Krishnasamy, B. Gomathi, "Task scheduling algorithm based on hybrid particle swarm optimization in cloud computing environment," *J. Theor. Appl. Inf. Technol.* 55, 33–38 (2013).
- [50] A. Verma, S. Kaushal, "Bi-Criteria Priority based Particle Swarm Optimization workflow scheduling algorithm for cloud," In: *Engineering and Computational Sciences (RAECS)*, 2014 Recent Advances in, 2014, pp. 1–6.
- [51] Y. Xiaoguang, C. Tingbin, Z. Qisong, "Research on cloud computing schedule based on improved hybrid PSO," In: *Computer Science and Network Technology (ICCSNT)*, 2013 3rd International Conference on, 2013, pp. 388–391
- [52] P. Visalakshi, S. Sivanandam, "Dynamic task scheduling with load balancing using hybrid particle swarm optimization," *Int. J. Open Problems Compt. Math.* 2, 475–488 (2009).
- [53] S.-J. Xue, W. Wu, "Scheduling workflow in cloud computing based on hybrid particle swarm algorithm," *TELKOMNIKA Indones. J. Electr. Eng.* 10, 1560–1566 (2012).

- [54] E.S. Sharma, G. Kaur, "Optimized utilization of resources using improved particle swarm optimization based task scheduling algorithms in cloud computing," *Int. J. Emerg. Technol. Adv. Eng.* 4, 110–115 (2014).
- [55] J.U. JieHui, B.A.O. WeiZheng, W.A.N.G. ZhongYou, W.A.N.G. Ya, L.I. WenJuan, "Research for the task scheduling algorithm optimization based on hybrid PSO and ACO for cloud computing," *Int. J. Grid Distrib. Comput.* 7(87–96), 2014 (2014).
- [56] M.Z. Zahraa Tarek, F.A. Omara, "Pso optimization algorithm for task scheduling on the cloud computing environment," *Int. J. Comput. Technol.* 13 (2014).
- [57] S. Pandey, L. Wu, S.M. Guru, R. Buyya, "A particle swarm optimization-based heuristic for scheduling workflow applications in cloud computing environments," In: *Advanced Information Networking and Applications (AINA)*, 2010 24th IEEE International Conference on, 2010, pp. 400–407.
- [58] R. Pragaladan, R. Maheswari, "Improve workflow scheduling technique for novel particle swarm optimization in cloud environment," *Int. J. Eng. Res. Gen. Sci.* 2, 5 (2014).
- [59] S. Chitra, B. Madhusudhanan, G. Sakthidharan, P. Saravanan, "Local minima jump PSO for workflow scheduling in cloud computing environments," In: *Advanced in Computer Science and its Applications*, pp. 1225–1234. Springer (2014).
- [60] F. Ramezani J., Lu, F. Hussain, "Task scheduling optimization in cloud computing applying multiobjective particle swarm optimization," in *Service-Oriented Computing*, pp. 237–251. Springer (2013).
- [61] G. Zhao, "Cost-Aware Scheduling Algorithm Based on PSO in Cloud Computing Environment," *International Journal of Grid & Distributed Computing* 7, 33–42 (2014).
- [62] S.A.M. Solmaz Abdi, S. Sharifian, "Task scheduling using modified PSO algorithm in cloud computing environment," In: *International Conference on Machine Learning, Electrical and Mechanical Engineering (ICMLEME'2014)*, Dubai, 2014.
- [63] A. Ramathilagam, K. Vijayalakshmi, "A Survey of Scheduling Algorithm in Cloud Computing Environment," *International Science Press, IJCTA*, 9(36), 2016, pp. 137-145.
- [64] N. Sharma, S. Tyagi, S. Atri, "A Survey on Heuristic Approach for Task Scheduling in Cloud Computing," *International Journal of Advance Research in Computer Science*, vol. 8, pp. 1089-1092, March-April 2017.
- [65] S. Sindhu, S. Mukherjee, "Efficient Task Scheduling Algorithms for Cloud Computing Environment," Springer-Verlag Berlin Heidelberg, pp. 79-83, 2011.
- [66] U. Bhoi, P. N. Ramanuj, "Enhanced Max-min Task Scheduling Algorithm in Cloud Computing," *International Journal Of Application or Innovation in Engineering and Management(IJAIEM)*, vol. 2, no. 4, pp. 259-264, April 2013.
- [67] Shimpy, J. Sidhu, "Different Scheduling Algorithms in Different Cloud Environment," *International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE)*, vol. 3, no. 9, pp. 8003-8006, September 2014.
- [68] R. kaur, P. K. Patra, "Resource Allocation with improved MinMin Algorithm," *International Journal of Computer Applications(IJCA)*, vol. 76, pp. 61-67, August 2013.
- [69] R. J. S. Raj, S. V. M. Prasad, "Survey on Variants of Heuristic Algorithms for Scheduling Workflow of Tasks," *International Conference on Circuit, Power and Computing Technologies [ICCPCT]*, IEEE, 2016.
- [70] G. Pate, R. Mehta, "A Survey on Various Task Scheduling Algorithm in Cloud Computing," *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, vol. 3, no. 3, pp. 715-717, March 2014.
- [71] A. Razaque, N. R. Vennapusa, N. Soni, G. S. Janapati, k. R. Vangala, "Task Scheduling in Cloud Computing," in *Long Island Systems, Applications and Technology Conference (LISAT)*, 2016 IEEE, 2016.
- [72] S. Mittal, A. Katal, "An Optimized Task Scheduling Algorithm in Cloud," *IEEE 6th International Advanced Computing Conference*, pp. 197-202, 2016.
- [73] B. Santhosh, D. H. Manjaiah, "An Improved Task Scheduling Algorithm based on Max-min for Cloud Computing," *International Journal of Innovative Research in Computer and Communication Engineering(IJIRCCE)*, vol. 2, no. 2, pp. 84-88, May 2014.
- [74] K. Etminani, M. Naghibzadeh, "A Min-Min Max-Min Selective Algoritm for Grid Task Scheduling," in *ICI 2007. 3rd IEEE/IFIP International Conference in Central Asia*, 2007.
- [75] B. P. Rimal, M. Maier, "Workflow Scheduling in Multi-Tenant Cloud Computing Environments," *IEEE Transaction on Parallel and Distributed Systems*, pp. 1-14, 2015.
- [76] X. Wu, M. Deng, R. Zhang, B. Zeng, S. Zhou, "Task Scheduling Algorithm Based on QoS-driven in Cloud Computing," Elsevier publications, pp. 1162-1169, 2013.
- [77] N. Bansal, A. Awasthi, S. Bansal, "Task Scheduling Algorithms with Multiple Factor in Cloud Computing Environment," Springer India, pp. 619-627, 2016.
- [78] P. N. Nimbalkar, "A Survey on Opportunistic Scheduling for Traffic Networks with Multiple Constraints," *International Journal of Innovative Research in Computer and Communication Engineering*, Vol. 5, Issue 5, May 2017. Website: [www.ijircce.com](http://www.ijircce.com).
- [79] Kulkarni, S. Suresh, C. Rosenberg, "Opportunistic scheduling: generalizations to include multiple constraints, multiple interfaces, and short term fairness," *Wireless Networks* 11.5 (2005): 557-569.
- [80] Neely, J. Michael, "Opportunistic scheduling with worst case delay guarantees in single and multi-hop networks," *INFOCOM*, 2011 Proceedings IEEE, 2011.
- [81] Khattab, KF Ahmed, K. MF Elsayed, "Opportunistic scheduling of delay sensitive traffic in OFDMA-based wireless network," proceedings of the 2006 International Symposium on World of Wireless, Mobile and Multimedia Networks. IEEE Computer Society, 2006.
- [82] Liu, Xin, E. KP Chong, N. B. Shroff, "Optimal opportunistic scheduling in wireless networks". *Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th*. Vol. 3. IEEE, 2003.
- [83] N. Abbas, Y. Zhang, A. Taherkordi, T. Skeie, "Mobile edge computing: A survey," *IEEE Internet of Things*, doi:10.1109/IJOT.2017.2750180.
- [84] G. Orsini, D. Bade, and W. Lamersdorf, "Computing at the mobile edge: Designing elastic android applications for computation offloading," in 2015 8th IFIP Wireless and Mobile Networking Conference (WMNC), Oct 2015, pp. 112–119.
- [85] E. Borgia, R. Bruno, M. Conti, D. Mascitti, A. Passarella, "Mobile edge clouds for information-centric iot services," in 2016 IEEE Symposium on Computers and Communication (ISCC), June 2016, pp. 422–428.
- [86] K. Kai, W. Cong, L. Tao, "Fog computing for vehicular ad-hoc networks: paradigms, scenarios, and issues," *The Journal of China Universities of Posts and Telecommunications*, vol. 23, no. 2, pp. 56–96, 2016.
- [87] R. Roman, J. Lopez, M. Mambo, "Mobile edge computing, fog *et al.*: A survey and analysis of security threats and challenges," *Future Generation Computer Systems*, pp. –, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0167739X16305635>

- [88] A. Ahmed E. Ahmed, "A survey on mobile edge computing," in 2016 10th International Conference on Intelligent Systems and Control (ISCO), 2016, pp. 1–8.
- [89] D. Kovachev, T. Yu, R. Klamma, "Adaptive computation offloading from mobile devices into the cloud", in 2012 IEEE 10th International Symposium on Parallel and Distributed Processing with Applications, July 2012, pp. 784–791.
- [90] S. Sardellitti, G. Scutari, S. Barbarossa, "Joint optimization of radio and computational resources for multicell mobile-edge computing," IEEE Transactions on Signal and Information Processing over Networks, vol. 1, no. 2, pp. 89–103, June 2015.
- [91] N. Takahashi, H. Tanaka, R. Kawamura, "Analysis of process assignment in multi-tier mobile cloud computing and application to edge accelerated web browsing," in Mobile Cloud Computing, Services, and Engineering (MobileCloud), 2015 3rd IEEE International Conference on, March 2015, pp. 233–234.
- [92] K. Kumar, J. Liu, Y.-H. Lu, B. Bhargava, "A survey of computation offloading for mobile systems," Mobile Networks and Applications, vol. 18, no. 1, pp. 129–140, 2013. [Online]. Available: <http://dx.doi.org/10.1007/s11036-012-0368-0>
- [93] K. Habak, M. Ammar, K. A. Harras, E. Zegura, "Femto clouds: Leveraging mobile devices to provide cloud service at the edge," in 2015 IEEE 8th International Conference on Cloud Computing, June 2015, pp. 9–16.
- [94] X. Chen, L. Jiao, W. Li, X. Fu, "Efficient multi-user computation offloading for mobile-edge cloud computing," IEEE/ACM Transactions on Networking, vol. 24, no. 5, pp. 2795–2808, Oct 2016.
- [95] K. Zhang, Y. Mao, S. Leng, A. Vinel, Y. Zhang, "Delay constrained offloading for mobile edge computing in cloud-enabled vehicular networks," in 2016 8th International Workshop on Resilient Networks Design and Modeling (RNDM), Sept 2016, pp. 288–294.
- [96] S. Abdelwahab, B. Hamdaoui, M. Guizani, T. Znati, "Replisom: Disciplined tiny memory replication for massive iot devices in lte edge cloud," IEEE Internet of Things Journal, vol. 3, no. 3, pp. 327–338, June 2016.
- [97] S. Nunna, A. Kousaridas, M. Ibrahim, M. Dillinger, C. Thuemmler, H. Feussner, A. Schneider, "Enabling real-time context-aware collaboration through 5g and mobile edge computing," in Information Technology - New Generations (ITNG), 2015 12th International Conference on, April 2015, pp. 601–605.
- [98] Y. Jararweh, A. Doulat, A. Darabsch, M. Alsmirat, M. Al-Ayyoub, E. Benkhelifa, "Sdmec: Software defined system for mobile edge computing," in 2016 IEEE International Conference on Cloud Engineering Workshop (IC2EW), April 2016, pp. 88–93.
- [99] N. Kumar, S. Zeckaly, J. J. P. C. Rodrigues, "Vehicular delay tolerant networks for smart grid data management using mobile edge computing," IEEE Communications Magazine, vol. 54, no. 10, pp. 60–66, October 2016.
- [100] A.E.H.G. El-Barbary, L.A.A. El-Sayed, H.H. Aly, M.N. El-Derini, "A cloudlet architecture using mobile devices," in 2015 IEEE/ACS 12th International Conference of Computer Systems and Applications (AICCSA), Nov 2015, pp. 1–8.
- [101] M. T. Beck, S. Feld, A. Fichtner, C. Linnhoff-Popien, T. Schimper, "Me-volte: Network functions for energy-efficient video transcoding at the mobile edge," in Intelligence in Next Generation Networks (ICIN), 2015 18th International Conference on, Feb 2015, pp. 38–44.
- [102] F. Jalali, K. Hinton, R. Ayre, T. Alpcan, R. S. Tucker, "Fog computing may help to save energy in cloud computing," IEEE Journal on Selected Areas in Communications, vol. 34, no. 5, pp. 1728–1739, May 2016.
- [103] W. Gao, "Opportunistic peer-to-peer mobile cloud computing at the tactical edge," in 2014 IEEE Military Communications Conference, Oct 2014, pp. 1614–1620.
- [104] Akashdeep, Karanjeet S. Kahlon, Harish Kumar, "Survey of Scheduling Algorithm in IEEE 802.16PMP Network," Egyptian Informatics Journal vol.15, 25-36, 2014.
- [105] A. Janiaka,W. A. Janiakb,T. Krysiaka,T. Kwiatkowski, "A survey on scheduling problems with due windows," European Journal of Operational Research 242(2015), 347–357.
- [106] S. Yi, Z. Qin, and Q. Li, "Security and privacy issues of fog computing: A survey," in International Conference on Wireless Algorithms, Systems and Applications (WASA), 2015.
- [107] KA Deshmame and BT Pandhare, "Survey on scheduling algorithms in cloud computing," International Journal of Emerging Trend in Engineering and Basic Sciences, 2(1):2102015, 2015.
- [108] L. Xiaojun, E. K. P. Chong, N. B. Shroff, "Opportunistic transmission scheduling with resource-sharing constraints in wireless networks," IEEE Journal on Selected Areas in Communications 19.10 (2001): 2053–2064.
- [109] S. Singh, I. Chana, "A survey on resource scheduling in cloud computing: issues and challenges," J. Grid Computing. 14 (2) (2016) 217–264. <http://dx.doi.org/10.1007/s10723-015-9359-2>.
- [110] M. Kaur, "Survey of Task Scheduling Algorithms in Cloud Computing," International Journal of Innovative Research in Computer and Communication Engineering, Vol. 5, Issue 4, April 2017. Website: [www.ijircce.com](http://www.ijircce.com).
- [111] M. A. Adibi, M. Zandieh, M. Amiri, "Multi-objective scheduling of dynamic job shop using variable neighborhood search," Expert Systems with Applications, 37(1), 282–287, 2010.
- [112] M. Hegde, P. Kumar, K. R. Vasudev, N. N. Sowmya, S. V. R. Anand, A. Kumar, J. Kuri, "Experiences with a Centralized Scheduling Approach for Performance Management of IEEE 802.11 Wireless LANs," IEEE/ACM Trans. Netw. 21 (2) (2013) 648–662.
- [113] K. Karanasos, S. Rao, C. Curino, C. Douglas, K. Chaliparambil, G. M. Fumarola, S. Heddaya, R. Ramakrishnan, S. Sakalanaga, "Mercury: Hybrid centralized and distributed scheduling in large shared clusters," In Proceedings of the 2015 USENIX Conference on Usenix Annual Technical Conference (Berkeley, CA, USA, 2015), USENIX ATC '15, USENIX Association, pp. 485–497.
- [114] W. Niu, Y. Xu, J. Wu, Y. Tan (2015) "A Centralized Scheduling Approach to Multi-Agent Coordination," In Proceedings of the 21st International Conference on Industrial Engineering and Engineering Management 2014, pp. 171–175.
- [115] D. Fu, J. Yang, X. Ling, H. Zhang, "Load-aware hybrid scheduling in large compute clusters," IEEE Symposium on Computers and Communication (ISCC), 2016.
- [116] O. D. Ramos-Cantor, J. Belschner, G. Hegde, M. Pesavento, "Centralized Coordinated Scheduling in LTE-Advanced Networks," arXiv:1702.03890v1, [cs.NI] submitted 13 Feb 2017.
- [117] J. Rintanen, "Scheduling with Contingent Resources and Tasks," Twenty-Third International Conference on Automated Planning and Scheduling, Rome, Italy, June 2013.
- [\*118] L. Zhou, Z. Yang, Y. Wen, and J. J. P. C. Rodrigues, "Distributed wireless video scheduling with delayed control information," IEEE Trans. Circuits Syst. Video Technol., vol. 24, no. 5, pp. 889–901, May 2014.
- [119] A. Hamzadayi, G. Yildiz, "Modeling and solving static m identical parallel machines scheduling problem with a common server and sequence dependent setup times," Computers & Industrial Engineering, 106 (2017), pages 287–298.
- [120] M. Skutella, M. Sviridenko, M. Uetz, "Unrelated machine scheduling with stochastic processing times," Math. Oper. Res., 41(3):851–864, 2016.

- [121] F. A. Chudak, "A min-sum 3/2-approximation algorithm for scheduling unrelated parallel machines," *Journal of Scheduling* 2 73–77, 1999.
- [122] A. S. Schulz, M. Skutella, "Scheduling unrelated machines by randomized rounding," *SIAM Journal on Discrete Mathematics* 15 450–469, 2002.
- [123] Mohring, R. H., A. S. Schulz, M. Uetz, "Approximation in stochastic scheduling: The power of LP-based priority policies," *Journal of the ACM* 46 924–942, 1999.
- [124] J. Sethuraman, M. S. Squillante, "Optimal scheduling of multiclass parallel machines," *Proceedings of the 10th Annual ACM-SIAM Symposium on Discrete Algorithms*. 963–964, 1999.
- [125] M. Skutella, "Convex quadratic and semidefinite programming relaxations in scheduling," *Journal of the ACM* 48 206–242, 2001.
- [126] P. Schuurman, G. J. Woeginger, "Polynomial time approximation algorithms for machine scheduling: Ten open problems," *Journal of Scheduling* 2 203–213, 1999.
- [127] E.O. Oyetunji, "Some Common Performance Measures in Scheduling Problems: Review Article," *Research Journal of Applied Sciences, Engineering and Technology* 1(2): 6-9, 2009.
- [128] D. M. Abdelkader, F. Omara, "Dynamic task scheduling algorithm with load balancing for heterogeneous computing system," *Egyptian Informatics Journal* (2012) 13, 135–145. www.elsevier.com/locate/eij, www.sciencedirect.com
- [129] M. Liu, F. Zheng, C. Chu, Y. Xu, "Optimal algorithms for online scheduling on parallel machines to minimize the makespan with a periodic availability constraint," *Theoretical Computer Science*, Vol. 412, p. 5225-5231, 2011.
- [130] D. Xu, Z. Cheng, Y. Yin, H. Li, "Makespan minimization for two parallel machines scheduling with a periodic availability constraint," *Computers & Operations Research*, Vol. 36, p. 1809-1812, 2009.
- [131] T. C. E. Chang, C. J. Hsu, D. L. Yang, "Unrelated parallel-machine scheduling with deteriorating maintenance activities," *Computers & Industrial Engineering* Vol. 60, p. 602-605, 2011.
- [132] S. Jeng, C. J. Hsu, D. L. Yang, "Unrelated parallel-machine scheduling with deteriorating maintenance activities," *Computers & Industrial Engineering*, vol. 62, pp. 1141-1143, 2012.
- [133] D. L. Yang, T.C.E Cheng, S. J. Yang, C. J. Hsu, "Unrelated parallel-machine scheduling with aging effects and multi-maintenance activities," *Computers & Operations Research*, Vo. 39, p. 1458-1464, 2012.
- [134] S. J. Yang, "Unrelated parallel-machine scheduling with deterioration effects and deteriorating multi-maintenance activities for minimizing the total completion time," *Applied Mathematical Modelling*, vol. 37-5, pp. 2995-3005, 2013.
- [135] S. Gawiejnowicz, A. Kononov, "Complexity and approximability of scheduling resumable proportionally deteriorating jobs," *European Journal of Operational Research* 200, 305–308, 2010.
- [136] D. Dolev, M. K. Warmuth, Scheduling precedence graphs of bounded height, *Journal of Algorithms*, 1984 - Elsevier.
- [137] G. Anastasi, M. Conti, M. Di Francesco, "Extending the Lifetime of Wireless Sensor Networks through Adaptive Sleep," *IEEE Trans. Industrial Informatics*, vol.5, no.3, pp. 351-365, 2009.
- [138] C.-W.Tsai, J.J. Rodrigues: "Metaheuristic scheduling for cloud: A survey," *IEEE Syst. J.* 8, 279–291, 2014.
- [139] Y. Wang, J. Wang, C. Wang, X. Song, "Research on resource scheduling of cloud based on improved particle swarm optimization algorithm," In *Advances in Brain Inspired Cognitive Systems*, pp. 118–125, Springer, 2013.
- [140] H. Kushner, P. Whiting, "Convergence of Proportional-Fair Sharing Algorithms Under General Conditions," *IEEE Transaction on Wireless Communication*, vol. 3, 4, 2004.
- [141] L. Yonghe, E. Knightly, "Opportunistic fair scheduling over multiple wireless channels," *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*. Vol. 2. IEEE, 2003.
- [142] L. Xin, E. K.P Chong, N. B. Shroff, "A framework for opportunistic scheduling in wireless networks," *Computer networks* 41.4 (2003): 451-474.
- [143] G. Konstantinos *et al.*, "Opportunistic scheduling of control tasks over shared wireless channels," *Cyber-Physical Systems (ICCP)*, 2014 ACM/IEEE International Conference on. IEEE, 2014.
- [144] A. A. Bhorkar *et al.*, "Adaptive opportunistic routing for wireless ad hoc networks," *IEEE/ACM Transactions On Networking* 20.1 (2012): 243-256.
- [145] Rasool, Jawad, *et al.*, "Opportunistic scheduling policies for improved throughput guarantees in wireless networks," *EURASIP Journal on Wireless Communications and Networking* 2011.1 (2011): 43.
- [146] L. Shihuan, L. Ying, R. Srikant, "Throughput-optimal opportunistic scheduling in the presence of flow-level dynamics," *IEEE/ACM Transactions on Networking* 19.4 (2011): 1057-1070.
- [147] C. Coskun, E.W. Knightly, "Opportunistic traffic scheduling over multiple network paths," *INFOCOM 2004. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*. Vol. 3. IEEE, 2004.
- [148] M. Wenguang, X. Wang, S. Wu, "Distributed Opportunistic Scheduling With QoS Constraints for Wireless Networks With Hybrid Links," *IEEE Transactions on Vehicular Technology* 65.10 (2016): 8511-8527.
- [149] T. Sheu-Sheu *et al.*, "Distributed opportunistic scheduling for ad-hoc communications under delay constraints," *INFOCOM, 2010 Proceedings IEEE*. IEEE, 2010.
- [150] Garcia-Saavedra, Andres, *et al.*, "Distributed opportunistic scheduling: A control theoretic approach," *Proceeding INFOCOM, IEEE*. 2012.
- [151] W. S. Herroelen, R. Leus, "Project scheduling under uncertainty: Survey and research potentials", *European Journal of Operational Research*, 165(2), 289-306. 2005.
- [152] S.B. Shaw, AK.Singh, "A survey on scheduling and load balancing techniques in cloud computing environment," *International Conference on Computer and Communication Technology (ICCCT)*, pp.87-95, 26-28 Sept. 2014.
- [153] J. A. Ruiz-Vanoye, O. Díaz-Parra, "Similarities between meta-heuristics algorithms and the science of life," *Cent Eur J Oper Res* (2011) 19: 445.
- [154] T. Lantharthong, N. Rugthaicharoencheep, "Network Reconfiguration for Load Balancing in Distribution System with Distributed Generation and Capacitor Placement," *World Academy of Science, Engineering and Technology International Journal of Electrical and Computer Engineering*, Vol:6, No:4, 2012.
- [155] B. Kaliski, "Polynomial Time," In: van Tilborg H.C.A. (eds) *Encyclopedia of Cryptography and Security*. Springer, Boston, MA, 2005.
- [156] C. H. Papadimitriou, Computational complexity. Reading, Mass.: Addison-Wesley, 1994, ISBN 0-201-53082-1.

- [157] K. Merhoum , M. Djeghaba, "Algorithme génétique pour le problème d'ordonnancement de type job-shop," 4th International Conference on Computer Integrated Manufacturing, CIP'2007 Sétif, Algeria, 03-04 November 2007.
- [158] H. B. Othman, Sur l'ordonnancement d'ateliers job-shop flexibles et flow-shop en industries pharmaceutiques Optimisation par algorithmes génétiques et essaims particulaires, Thèse de Doctorat délivré conjointement par l'École Centrale de Lille et l'École Nationale d'Ingénieurs de Tunis, Juillet 2009.

# Digitalization and Education System: A Survey

Saiful Islam

Lecturer at dept. of CSE

Daffodil International University

Dhaka, Bangladesh

E-mail: saiful.cse@diu.edu.bd

Nusrat Jahan

Dept. of CSE

Daffodil International University

Dhaka, Bangladesh

E-mail: nusratjahan.cse@diu.edu.bd

**Abstract**— Education is a basic need for every human being and digital education is the current trend and necessity for every students or learners to be more focused in their learning. In this paper authors worked with these current phenomena. Digital education helps students or learners to gather knowledge in easier and different ways than before. It also reduces the learning time. In traditional education system we were mostly dependent on text book or in instructor's speech. But nowadays it is easier to find any text book or any other learning materials by using digital educational tools. Another charismatic change that transformed human life is social networking. In terms of digital education social networks contribute a good portion of education. Among social networking services, Facebook has become most popular for communication with familiar and unfamiliar persons. The impact of the use of Facebook on students is very impactful. In this paper authors conducted a survey on various students for understanding the digitalization effect on educational purpose. Machine learning was applied for classified the happy and unhappy student with digitalization where focused time spent on educational purposes. Finally authors provided an analytical summary of digitalization effect on education based on their survey.

**Keywords**—Digitalization, E-learning, Machine learning, Social networking.

## I. INTRODUCTION

The history of Internet began with the development of electronic computers in 1950s. Like many other developed and developing countries, Internet in Bangladesh has also witnessed phenomenal growth. Although facing many constraints in expanding internet access and uses, development of the internet and information technology is one of the government's high priorities. In present era, social networking websites, such as MySpace and Facebook have been attracting a large number of participants. For example- in 2013, Internet users in Bangladesh increased to 33 million [1].

In social networks, each node represents a participant and each link between participants corresponds to real-world interactions or online interactions between them. One participant can give a trust value to another based on their interactions [2]. Online social networking sites have been attracting a large number of participants and are being used as the means for a variety of rich activities. For example, participants carry out business, and share photos and movies on the first generation (e.g., ebay.com) and second generation (e.g., facebook.com) social networking sites respectively [3]. Authors tried to find out the percentage of participants that

actually uses these social networking sites for digital education purpose. They also tried to determine is these sites really helping our student participants to gain more quicker and clear knowledge to grow as a learner or are they just wasting their time leading them to a poor career.

This paper is organized as follows: Section II gives a brief historical idea about digitalization's activity. Section III for a small overview of this paper work. In section IV discuss about result analysis. Section V includes the conclusion with future plan.

## II. RELATED WORK

In this section noted some previous discussion about digitalization, social networking and machine learning with digital education.

Umamaheswari. k and S. Niraimathi worked with student's socio-demographic variables like age, gender, name, class grade, proficiency and extra skill. Their data analysis result helps recruitment process on interview board through student's grade [7]. Sunday Tunmibi, Ayooluwa Arebesola, Pascal Adejobi, and Olaniyi Ibrahim was discussed the impact of e-learning and digitalization in primary and secondary school levels. They showed that majority of teachers agreed about e-learning helps our students to gather more knowledge and resources [4]. Manoj Kumar discussed on smart phone uses in education technology. He was also worked with the application of technical and professional studies in Indian education [5]. Pooja Thakar, Anil Mehta, and Manisha discussed about educational data mining which was based on different survey results. Machine learning helps us to find out informative information to solve a problem [6]. Radhika R Halde was introduced machine learning approach for predicting the student's performances and also compared different machine learning algorithms [12].

One of the digitalization impact of Bangladesh is 1 to class 10 including teacher's training and other necessary books are available at this website [www.ebook.gov.bd](http://www.ebook.gov.bd) website. The government provides laptops and multimedia projectors to 20,500 public and private educational institutions to improve the classroom teaching-learning process [18].

There are many online social networking sites like Facebook, Twitter, MySpace, eBay etc. Among all of these sites, participants of Facebook are the highest. Facebook is an online social networking service. Its name stems from the colloquial name for the book given to students at the start of the academic year by some American university

administrations to help students get to know each other [8]. Facebook (as of 2012) has about 180 petabytes of data a year and grows by over half a petabytes every 24 hours [9].

After this study here author worked with digitalization impact on education and they discussed the positive and negative answer of student's. Finally author provided a result which was focused the reason of this survey.

### III. OVERVIEW

Social networking is an internet based medium that makes a way to communicate with friends, family, classmates, customers and clients. Social networking can occur for social purposes, business purposes or both through sites such as Facebook, Twitter, and LinkedIn etc. [13]. But nowadays social networking also used for educational purposes. Facebook is the most popular social networking site. Facebook was founded in February 2004 by Mark Zuckerberg with his college roommates. As of September 2012, Facebook has over one billion active users [14]. There are many Facebook pages and groups that help us to gather information in different reasons. Finally it helps us to increases our knowledge besides social communication.

Machine learning algorithms can be categorized under two main streams: supervised learners and unsupervised learners [16]. The program is trained with a pre-defined set of training examples, which then facilitate its ability to reach an accurate conclusion when given new data is the supervised learning and Unsupervised machine learning is the program where given a bunch of data and must find patterns and relationships therein [17]. In this paper author focused on supervised learning approaches.

Digital Technology for education is defined with any process where the teacher or learner uses digital equipment such as a personal computer, a Laptop, tablet, MP3 player, or console to access digital tools such as learning platforms and virtual learning environments (VLEs) to improve their knowledge and skills. The Learning with Digital Technology comprises of ICT products such as teleconferencing, email, audio, television lessons, radio broadcasts, interactive voice response system etc. [19]. Day by day all over the world go through with digitalization and our education system is one of the biggest fields where we can introduce more digitalization.

### IV. DATA ANALYSIS AND RESULT

Authors have conducted an online survey on the basis of Internet access among several students of different educational institutions throughout Bangladesh [10]. Here collected necessary real time data on the basis of some questions from students of corresponding regions. Here considered responses from 283 students. All of the participants use Internet at different time. In this paper author also introduced machine learning to classify the student's based on the two fields-Gender and Happy/Unhappy with digitalization (H/UH). This paper also realized the male and female student's interest on digitalization. In this section discussed five issue related with the survey and those actually focused on digital education effect on education.

#### A. Effectiveness

Although the traditional education in our country i.e. writing on the board with chalk or white board marker is still preferable to many of the teachers, our student's thoughts are otherwise. On our survey one of the most important questions was the thought of our students about the effectiveness of Digital Education. A maximum of number of 95.4% student are happy while the teachers uses digitization tools inside the classroom. Only 4.6% students stayed with the unhappy group. As our survey result is suggesting most of our students are changing and loving digitization so our teachers should keep that fact in mind and prepare their content based on that.

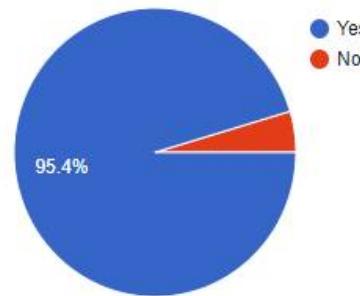


Figure 1. Effectiveness of digitalization.

Another aspect of the effect of digitization tools inside the classroom was found when author raised a question in their survey that is “Do you think these digitized tools are effecting your classroom studying?” the answer was varying with the happy unhappy ratio. Author found that 83.4% agrees with the positive effect of digitalization inside the classroom while 16.6% of students said that traditional system is not good or bad then the digitalized education. In this section author found that students have different thinking about digitalization on classroom. Some students have no idea about practices of digital tools for educational purposes.

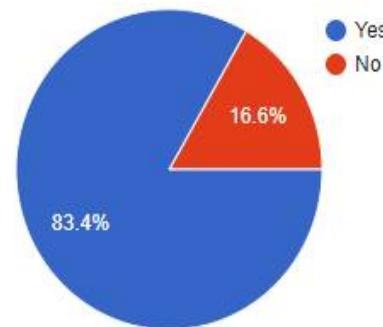


Figure 2. Digitalization effect on classroom study.

#### B. Social Network

After the vision 2021 of current Bangladesh Govt. in 2008 of a digital world many of our pupil is now online and a large portion of this online activity is based on social networks, mostly Facebook, Twitter, WhatsApp and many others. Nowadays social networks contribute a good portion of Internet traffic and thus attract tremendous research interests. Our students carry a large participation in this activity. In this survey result we found 95.3% students are fond of Facebook and 20.4% of twitter in their social networking activities.

Only a mere of 1.1% are fond of YouTube, while there are some other social network users in our survey. We also found that students are spending more than 2-3 hrs for social networking on daily basis.

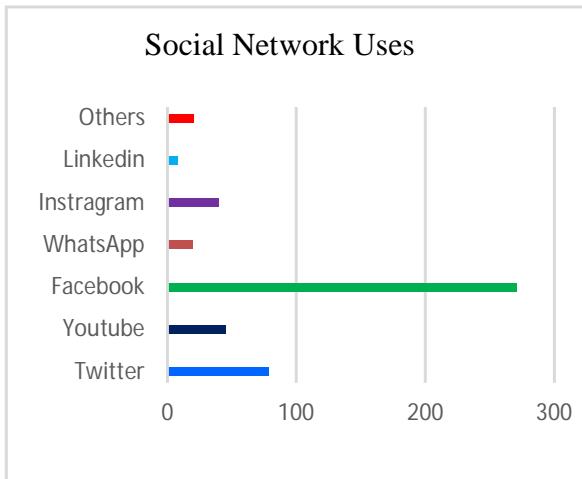


Figure 3. Social Networking.

#### C. Gender

Since 1971 population of Bangladesh is rising in an alarming rate. According to the population structure in Bangladesh male percentage per 100 female of an age of 20-24 is 76.8% [11]. But as our survey result shows that female student percentage is less than 20%. It is very concerning fact for our women education and development because in our country most of the family's children are growing up in a mother centric environment. And it is still believed that "An educated mother can provide an educated nation". In this survey we found that the percentage of female student's is 18.4% that is not too good in our nation.

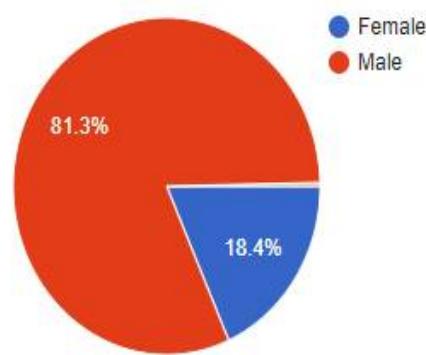


Figure 4. Percentage of Gender.

#### D. Education Time

Our survey was based on digital education and its effect, where one of the key question was "Per day how much time do you spend for education by using Internet?" The result of this question is so impressive that the average spending time for education purpose is more than 3hrs. Most of our students

like to spend this time on Google, YouTube, Wikipedia and others. In table we show the survey result for this scenario.

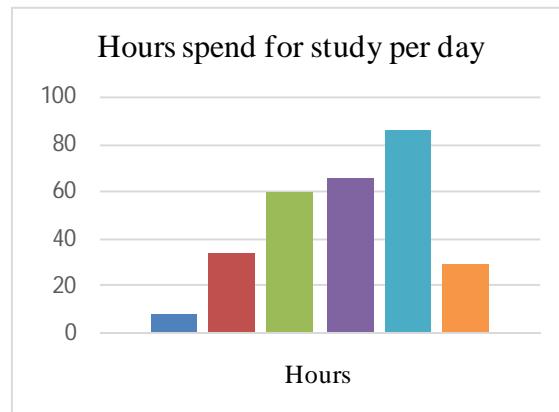


Figure 5. Hours spend for educational purposes per day.

#### E. Helpful tools inside class room

In this digital era student like to learn in their classroom by using power point presentation slide, YouTube, Google classroom and few others. Based on our survey result we found that 85.4% students like to use these digital tools inside classroom.

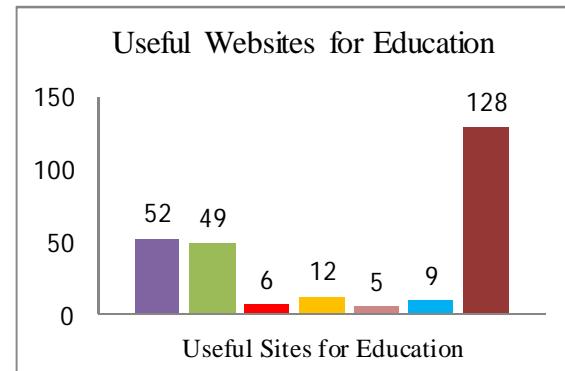


Figure 6. Useful website for education.

#### F. Machine Learning and Survey

In this section author was used Weka 3.8.0 for understood the male and female students ration about happy with digitalization effect and also here we found that the ration of male and female in our graduation level of education. In Figure 7 represent the decision tree.

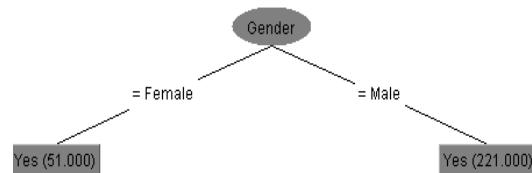


Figure 7. No. of Happy participant with digitalization.

Table 1 for presented the total discussion result at a glance. Here focused the above five terminology- Effectiveness, Social Network, Gender, Education Time and Helpful website and tools inside class room.

TABLE I. SURVEY RESULT.

Field	Survey Result	
Effectiveness	Digitalization	Y-95.4% N-4.6%
	Digitalization on classroom	Y-83.4% N-16.6%
Social Network	Facebook	
Gender	Male-81.3% Female-18.4%	
Education Time	Average 3Hrs/Day	
Helpful website & tools	Google.com Youtube.com Power point presentation	

## V. CONCLUSION

Today most of the learning styles have been converted into digital education system. Digital education also extends through social networks. Here the survey result showed the effect of social networking in our education site. So in future we can work with e-learning through Facebook that will more interactive and easy to access. If this study attract the concerned authority, then it will be helpful for all of the learners who use Facebook across the world. In future, authors have a target to work with machine learning approaches where they consider digital education system.

## REFERENCES

- [1]. Rafiqul Islam Azad, "33 million Internet users in Bangladesh". The Independent [Dhaka: Independent Publications Limited, 31 May 2013].
- [2]. Guanfeng Liu, Yan Wang, and Mehmet A. Orgun, Ee-Peng Lim, "A Heuristic Algorithm for Trust-Oriented Service Provider Selection in Complex Social Networks" 2010 IEEE.
- [3]. G. Liu, Y. Wang, and M. Orgun, "Trust inference in complex trust-oriented social networks". In IEEE CSE'09, pages 996–1001.
- [4]. Sunday Tunmibi, Ayooluwa Aregbesola, Pascal Adejobi, and Olaniyi Ibrahim, "Impact of E-Learning and Digitalization in Primary and Secondary Schools". Journal of Education and Practice, Vol.6, No.17, 2015.
- [5]. Manoj Kumar, "Impact of the Evolution of Smart Phones in Education Technology and its Application in Technical and Professional Studies: Indian Perspective". International Journal of Managing Information Technology (IJMIT), Vol.3, No.3, August 2011.
- [6]. Pooja Thakar, Anil Mehta, and Manisha, "Performance Analysis and Prediction in Educational Data Mining: A Research Travelogue". International Journal of Computer Applications (0975 – 8887) Volume 110 – No. 15, January 2015.
- [7]. Umamaheswari. K , and S. Niraimathi, "A Study on Student Data Analysis Using Data Mining Techniques". International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 8, August 2013.

- [8]. Eldon, Eric, "2008 Growth Puts Facebook In Better Position to Make Money". [Venture Beat (San Francisco), December 18, 2008].
- [9]. Sharwood, and Simon, "Facebook warehousing 180 PETABYTES of data a year". [9th November 2012].
- [10]. <https://docs.google.com/forms/d/e/1FAIpQLSe03grxoq5OdBuqzW0-mrOoNdm-eVzX3leYATnbCLJP0-Zhw/viewform>
- [11]. [https://en.wikipedia.org/wiki/Demographics\\_of\\_Bangladesh](https://en.wikipedia.org/wiki/Demographics_of_Bangladesh)
- [12]. Radhika R Halde, "Application of Machine Learning algorithms for betterment in Education system." 2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT).
- [13]. <https://www.investopedia.com/terms/s/social-networking.asp>
- [14]. "Facebook Tops Billion-User Mark". [The Wall Street Journal (New York). October 4, 2012].
- [15]. Peña-Ayala, and Alejandro, "Educational data mining: A survey and a data mining-based analysis of recent works". Expert systems with applications 41.4 (2014): 1432-1462.
- [16]. Kotsiantis, Sotiris B., and Panayiotis E. Pintelas, "Predicting student'smarks in Hellenic open university". In Fifth IEEE internationalconference on Advanced learning technologies, ICALT 2005, pp. 664-668. IEEE, 2005.
- [17]. <https://www.toptal.com/machine-learning/machine-learning-theory-an-introductory-primer>
- [18]. The Ministry of Education website. Available:[http://www.moedu.gov.bd/index.php?option=com\\_docman&task=cat\\_view&gid=161&dir=DESC&order=name&Itemid=9999999&limit=10&limitstart=0](http://www.moedu.gov.bd/index.php?option=com_docman&task=cat_view&gid=161&dir=DESC&order=name&Itemid=9999999&limit=10&limitstart=0), 2009.
- [19]. Bikas C. Sanyal, "New functions of higher education and ICT to achieve education for all". Paper prepared for the Expert Roundtable on University and Technology-for- Literacy and Education Partnership in Developing Countries, International Institute for Educational Planning, UNESCO, September 10 to 12, Paris 2001.

## AUTHORS PROFILE



### Saiful Islam

He completed his Graduation in CSE from Daffodil International University, Bangladesh and Master in Information Technology from University of Dhaka, IIT.

He has 3 years of teaching experience. His research interests are Data Analysis and Human Computer Interaction.



### Nusrat Jahan

She completed her Graduation in CSE from Patuakhali Science and Technology University, Bangladesh and Master in Information Technology from University of Dhaka, IIT.

She is having 3 years of teaching experience. Her research interests are machine learning, Data Analysis and image processing.

# Implementation On Data Security Approach In Dynamic Multi Hop Communication

Dr. S. A. Ubale

Zeal college of Engineering and  
research  
Narhe 41  
Pune, India  
[swapnaja.b.more@gmail.com](mailto:swapnaja.b.more@gmail.com)

Anuradha V. Khare

Zeal college of Engineering and  
research  
Narhe 41  
Pune, India  
[anuradkhare53@gmail.com](mailto:anuradkhare53@gmail.com)

Dr. Sulbha S. Apte

Walchand Institute of Technology  
Solapur 06  
Solapur, India

**Abstract**— In remote sensor arrange messages are exchanged between the different source and goal matches agreeably such way that multi-jump parcel transmission is utilized. These information bundles are exchanged from the middle of the road hub to sink hub by sending a parcel to goal hubs. Where each hub overhears transmission close neighbor hub. To dodge this we propose novel approach with proficient steering convention i.e. most brief way directing and conveyed hub steering calculation. Proposed work additionally concentrates on Automatic Repeat Request and Deterministic Network coding. We spread this work by the end to end message encoding instrument. To upgrade hub security match shrewd key era is utilized, in which combined conveying hub is allocated with combine key to making secure correspondence. End to end. We dissect both single and numerous hubs and look at basic ARQ and deterministic system coding as strategies for transmission.

**Keywords:** SINR, Mesh Network, Sensor Deployment.

## I. INTRODUCTION

In multi-jump remote system parcel transmission by safeguarding the privacy of transitional hubs, with the goal that information sent to a hub is not shared by some other hub. Additionally, in which secrecy is a bit much, it might be not secure to consider that hubs will dependably remain uncompromised. In remote system hubs, information secret can be seen as a security to stay away from a traded off hub from getting to data from other uncompromised hubs. In a multi-bounce organize, as information parcels are exchanged, middle of the road hubs gets all or part of the information bundle through straightforward transmission of system hub by means of multi-jump arrange mold, while exchanging classified messages. Proposed work alludes productive calculations for secret multiuser correspondence over multi-bounce remote systems. The metric we use to quantify the privacy is the shared data spillage rate to the transfer hubs, i.e., the equivocation rate. We require this rate to be self-assertively little with high likelihood and force this in the asset allotment issue by means of an extra limitation. We

consider down to earth postpone necessities for every client, which wipes out the likelihood of encoding over a discretionarilylong piece.

## II. PROBLEM STATEMENT

Proposed system present the problem of network utility maximization, into which confidentiality is incorporated as an additional quality of service constraint. Secure message transmission between the source and a destination node with less overhead cost. Data transfer using multi-hop with minimum overhead and secure communication among network node. Proposed system resolve problem of distributed scheduling. Cross-layer node allocation problem with confidentiality in a cellular wireless network, where users transmit information to the base station, confidentially from the other users.

## III. LITERATURE SURVEY

This system proposed private and public channels to minimize the use of the (more expensive) private channel in terms of the required level of security. This work considers both single and multiple users and compares simple ARQ and deterministic network coding as methods of transmission [1]. This paper design secure communications of one source-destination pair with the help of multiple cooperating intermediate nodes in the presence of one or more eavesdroppers. Three Cooperative schemes are considered: decode-and-forward (DF), amplify-and-forward (AF), and cooperative jamming (CJ). For these schemes, the relays transmit a weighted version of a re-encoded noise-free message signal (for DF), a received noisy source signal (for AF), or a common jamming signal (for CJ)[2]. This paper considers secure network coding with nonuniform or restricted wiretap sets, for example, networks with unequal link capacities where a wiretapper can wiretap any subset of links, or networks where only a subset of links can be wiretapped [3]. The scheme does not require eavesdropper CSI (only the statistical knowledge is assumed) and the secure

throughput per node increases as we add more legitimate users to the network in this setting. Finally, the effect of eavesdropper collusion on the performance of the proposed schemes is characterized [4]. We characterize the secrecy capacity in terms of generalized eigenvalues when the sender and eavesdropper have multiple antennas, the intended receiver has a single antenna, and the channel matrices are fixed and known to all the terminals and show that a beamforming strategy is capacity-achieving. In addition, we study a masked beam forming the scheme that radiates power isotropically in all directions and shows that it attains near-optimal performance in the high SNR regime [5].

#### IV. SYSTEM ARCHITECTURE

In existing hop to hop communication in wireless sensor network considered to succumb to the vulnerability of data transmission. Due to hop by hop communication increased cost for packet transmission, the existing system uses security mechanism as a node to node authentication among network resources. Hop to hop identity of intermediate node compromise security threats. To avoid security threat they use digital signature authentication at node level for communication or packet transmission. In the existing system, message transmission is done through all neighbors between the source and destination nodes, which result in overhearing and increase overhead between nodes. Also, it leads to compromised node communication in wireless sensor communication.

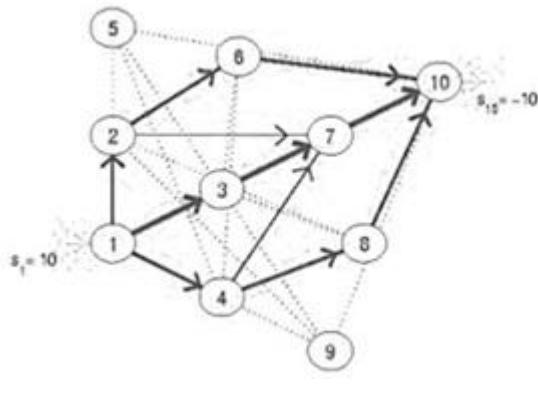


Figure. 1. Proposed System ( Architecture ) and Working.

Proposed system implements an optimal dynamic policy for the case in which the number of blocks across which secrecy encoding is performed is asymptotically large. Next, to that, This work propagate encoding between a finite number of data packets, which removes the possibility of achieving perfect secrecy. In this case, proposed work design a dynamic policy to select the encoding rates for every data packet, based on the instantaneous channel state information, queue states and secrecy humiliation requirements. By numerical analysis, we observe that the proposed scheme approaches the optimal rates asymptotically with increasing block size.

Finally, we address the consequences of practical implementation issues such as infrequent queue updates and de-centralized scheduling. Existing work present the efficiency of our policies by numerical studies under various network conditions. Next to this work proposed system contribute to deterministic network coding Automation of repeat packet request mechanism to actively transfer data packet. This help to network costs and other system parameters were just designed as constants in our work the network costs are related to physical layer parameters such as channel encoding parameters and transmission power. Here proposed system design in the way, which formulate problem by adding noise to original message or request at a destination.

The proposed system also formulate problem ARQ case in which automatic repeat request is sent between numbers of the time slot during packet sending. Where packets are generally transferred via the private channel and public channel from source to destination. These packets are generally geometrically distributed among network nodes.

#### V. ALGORITHM DETAILS

##### A. Generate an RSA key pair

**Input :** Required modulus bit length, k.

**Output :** An RSA key pair ((N,e), d) where N is the modulus, the product of two primes (N=pq) not exceeding k bits in length; e is the public exponent, a number less than and coprime to (p-1)(q-1); and d is the private exponent such that  $ed \equiv 1 \pmod{(p-1)(q-1)}$ .

```

Select a value of e from { 3, 5, 17, 257, 65537 }
repeat
    p ← genffiprime(k/2)
until (p mod e) ≠ 1
repeat
    q ← genffiprime(k - k/2)
until (q mod e) ≠ 1
N ← pq
L ← (p-1)(q-1)
d ← modffinv(e, L)
Return (N, e, d)

```

The system has classified into the different sets like below

$$\text{Sys} = \{\text{inp, process, out, analysis}\}$$

$$\text{Inp} = \{\text{D1, D2, ..., Dn}\}$$

That is the set of input data chunks

$$\text{EncData} \sum_{n=1}^m \text{Enc}(D) \dots \text{Enc}(Dn) \quad \dots \quad (1)$$

Equation (1) shows the data aggregation as well as data encryption process.

$$\text{Data} = \sum D[i] \quad \dots \quad (2)$$

Equation (2) shows the get the data from each node

$$\text{PlainData} \sum_{n=1}^m \text{Dec}(D_1) \dots \text{Dec}(D_n) \quad \dots \quad (3)$$

Equation (3) shows the data aggregation of cipher data on receiver phase with decryption process.

### B. Construction of updated BTC

**Input:** Initial source node sn, Destination node dn, Group of neighbor nodes nd [], each node id, each node energy eng.

**Output:** Source to destination path when data received success.

Step 1: User first selects the sn and dn

Step 2: choose the packet or file f for data transmission.

Step 3: if(f!=null) fd<= f

Step 4: read each byte b form fd when reach null

Step 5: send data; initialize cf1, cf2, pf1, pf2.

Step 6: while (nd[i] when reach NULL)

Cf1=nd[i].eng

Pf1= nd[i].id

Cf2=nd[i+1].eng

Pf2= nd[i+1].id

Step 7: if (cf1>cf2)

Cf2=null

Pf2=null

Else

Pf1=pf2

Cf1=cf2;

Pf2=null

Cf2=null

Step 8: end while

Step 9: repeat up to when reach at sink node.

## VI. EXPERIMENTAL SETUP

We run our experiments in NS2 simulator version 2.35 that has shown to produce realistic results. NS simulator runs TCL code, but here use both TCL and C++ code for header input. In our simulations, we use Infrastructure based network environment for communication. For providing access to the wireless network at any time used for the network selection.

WMN simulate in NS2.TCL file shows the simulation of all over architecture which proposed. For run.TCL use EvalVid Framework framework in NS2 simulator it also helps to store running connection information message using connection pattern file us1. NS2 trace file .tr can help to analyze results. It supports filtering, processing and displaying vector and scalar data. The results directory in the project folder contains us.tr file which is the files that store the performance results of the simulation. Based on us.tr file using the xgraph tool we execute graph of result parameters with

respect to x and y-axis parameters. Graphs files are of .awk extensions and are executable in the x-graph tool to plot the graph.

### A. Types of simulation

Parameter	Value
Simulator	Ns-allinone-2.35
Simulation Time	40sec
Channel Type	Wireless Channel
Propogation Model	Propogation Two Ray Ground
Medium	Phy/Wireless Phy
Standard	Mac/802 11
Logical Link Layer	LL
Antenna	Antenne/Omni Antenna
X dimension of the topography	1500
Y dimension of the topography	1000
Max Packet in ifq	1000
Adhoc Routing	AODV
Routing	DSR
Traffic	cbr

Table 1. Behaviour of parameters versus Simulation time for Different Nodes .

These Parameters are defined and evaluated below:

### B. Average End-to-End Delay

End-to-End Delay (E2ED) refers to time occupied by a data packet travel from a source to a destination in a network. Here only data that reaches successfully to the destination are considered. The minimum value of E2ED means the good performance of the protocol. The smallest amount value of end-to-end delay states superior performance of the protocol.

### C. Packet Delivery Ratio

The packet delivery ratio (PDR) defined as a ratio of numbers of data packets reached to target over the network to a number of packets generated. The greater amount value of packet delivery ratios states superior performance of the protocol.

### D. Throughput

Throughput can be defined as the ratio of the total bytes in data packets received by sink nodes to time from first packets generated at a source to the last packet received by sink nodes. The greater value of throughput states superior performance of the protocol.

### E. Energy Cosumption

Energy consumption is most important concepts in WSN. The lifetime of the sensor network is based on energy consumption

of the sensor node. Total energy consumption of the node defined as the difference between initial energy and final energy of the node. The smallest amount value of energy consumption states superior performance of the protocol.

#### F. For 100no of Nodes

##### 1) Delay versus Simulation Time

The end-to-end delay in SINGLE HOP, DUAL HOP and DDT with an increase in Simulation time. However, increasing treads in DUAL HOP and SINGLE HOP is much higher than Proposed as shown in Table 2. The smallest amount value of end-to-end delay states superior performance of the protocol. Figure 2 shows, the proposed system gives superior performance than other three protocols.

Delay				
Simulation Time	Multi Hop Proposed	Single Hop	Dual Hop	Distributed Data Transmission (DDT)
0.15	0.00562	0.00752	0.00622	0.00803
0.20	0.00578	0.00782	0.00653	0.00901

Table 2. Delay of 100 Nodes.

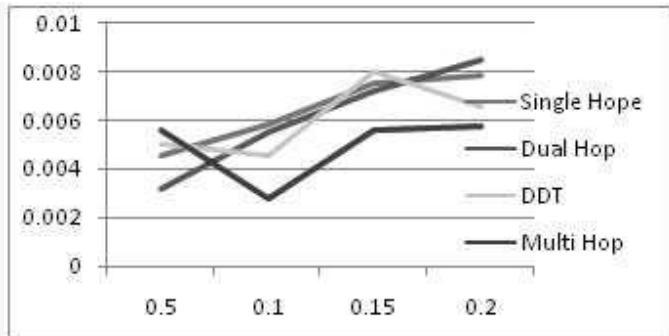


Fig. 2. Delay versus Simulation Time.

##### 2) Packet Delivery Ratio versus Simulation Time

The packet delivery ratio of SINGLE HOP, DUAL HOP, and DDT than proposed system decreases with increase in Simulation Time as shown in Table, However, decreasing treads in SINGLE HOP and DUAL HOP is much smaller than proposed approach. The greater amount value of packet delivery ratios states superior performance of the protocol as shown in Fig 3.

PDR				
Simulation Time	Multi Hop Proposed	Single Hop	Dual Hop	Distributed Data Transmission (DDT)
0.15	95.20	90.20	92.45	95.10
0.20	95.15	90.40	91.30	96.03

Table 3. PDR of 100 Nodes.

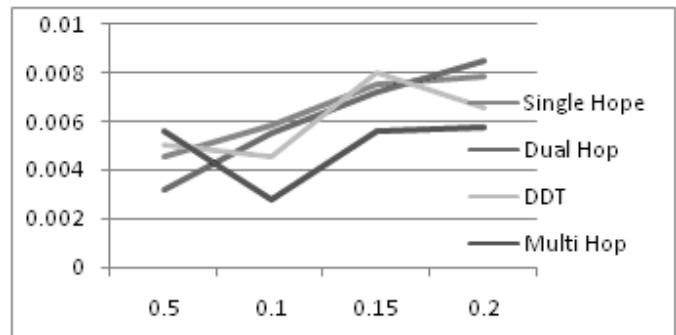


Fig. 3. PDR versus Simulation Time.

##### 3) Throughput versus Simulation time

Figure 4 shows the throughput under different networks scale in DUAL HOP, SINGLE HOP, DDT and Multi-Hop. The throughput in proposed, SINGLE HOP, DDT and DUAL HOP increases with increase in Simulation Time. The greater value of throughput states superior performance of the protocol as shown in Table 4.

Throughput				
Simulation Time	Multi Hop Proposed	Single Hop	Dual Hop	Distributed Data Transmission (DDT)
0.15	196.20	189.20	183.45	179.10
0.20	194.15	188.40	184.30	181.03

Table 4. Throughput of 100 Nodes.

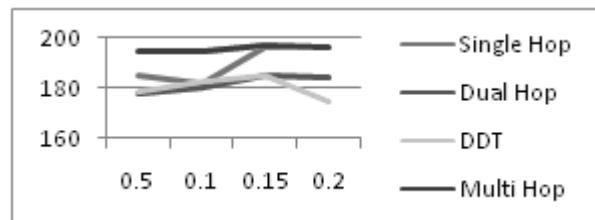


Fig. 4. Throughput versus Simulation Time.

##### 4) Energy versus Simulation time

The energy consumption of DUAL HOP , SINGLE HOP , DDT and Hybrid DUAL HOP decreases with increase in Simulation Time . However, decreasing treads in DUAL HOP and Proposed approach is much higher than SINGLE HOP , DDT as shown in Table 5. The smallest amount value

of energy consumption states superior performance of the protocol as shown in Fig 5.

Simulation Time	Energy			
	Multi Hop Proposed	Single Hop	Dual Hop	Distributed Data Transmission (DDT)
0.15	755	1120	1320	1760
0.20	956	1293.40	1570	1985

Table 5. Energy required for simulation of 100 Nodes.

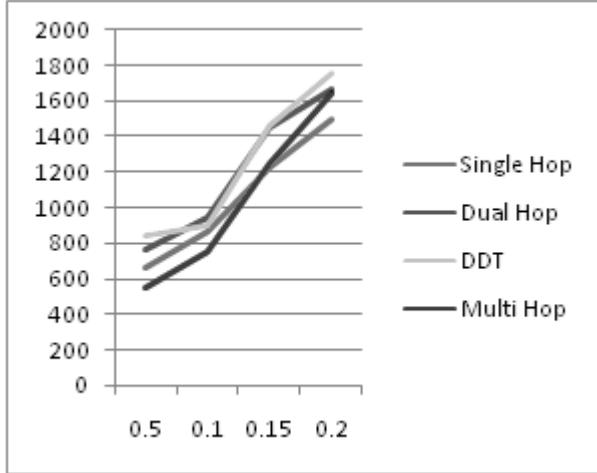


Fig. 5. Energy versus Simulation Time.

##### 5) Accuracy of System

In order to evaluate the performance of system performed. The network architecture considered is the following:

- A fixed base station (sink node) is located away from the sensor field.
- The sensor nodes are energy constrained with homogeneous initial energy allocation.
- Each sensor node senses the surroundings at a fixed rate and at all times its data to send to the base Station (data are sent if an event occurs).
- The sensor nodes are assumed to be stationary. However, the protocol can also support.

We compare the proposed system results with different existing system. Below table shows the comparison analysis of proposed system with some existing system

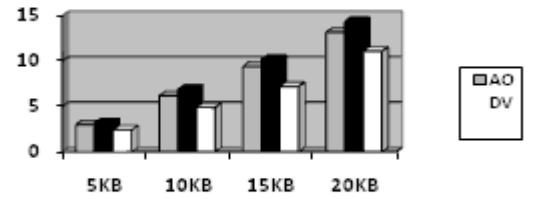


Fig. 6. System comparison proposed vs existing ( milliseconds ).

We consider energy evaluation for transmission which will conserve the node energy at the time of transmission, the system will select efficient path for communication with neighbor node at same time remaining network will sleep node.

## VII. CONCLUSION

Secure and effective way reproduction for parcel misfortunes and in addition directing progression. At the hub side, Pathfinder is an instrument which has a connection between's an arrangement of bundle ways and productively packs the way data utilizing way distinction. At the sink side, Pathfinder deduces parcel ways from the compacted data and utilizes astute way theory to reproduce the bundle ways with high remaking proportion.

Straightforward Automatic Repeat Request (ARQ), and Deterministic Network Coding (DNC), where in each vacancy the source shapes M directly autonomous deterministic blends of the M parcels and afterward utilize basic ARQ to transmit each straight mix dependably to the goal. We expect for this situation that the collector does not make an induction from the got straight blends but rather either disentangles the transmitted bundles or not.

## REFERENCES

- [1] N. Abuzainab and A. Ephremides, "Secure distributed information exchange", IEEE Trans. Inf. Theory, vol. 60, no. 2, pp. 1126–1135, Feb. [2014].
- [2] L. Dong, Z. Han, A. P. Petropulu, and H. V. Poor, "Improving wireless physical layer security via cooperating relays", IEEE Trans. Signal Process., vol. 58, no. 3, pp. 4033–4039, Mar. [2010].
- [3] T. Cui, T. Ho, and J. Kliewer, "On secure network coding with nonuniform or restricted wiretap sets", IEEE Trans. Inf. Theory, vol. 59, no. 1, pp. 166–176, Jan. [2013].
- [4] O. O. Koyluoglu, C. E. Koksal, and H. E. Gamal, "On secrecy capacity scaling in wireless networks", IEEE Trans. Inf. Theory, vol. 58, no. 5, pp. 3000–3015, May [2012].
- [5] A. Khisti and G. W. Wornell, "Secure transmissions with multiple antennas: Themisome wiretap channel", IEEE Trans. Inf. Theory, vol. 56, no. 7, pp. 3088–3014, July [2010].
- [6] C. E. Koksal, O. Ercetin, and Y. Sarikaya, "Control of wireless networks with secrecy", IEEE/ACM Trans. Netw., vol. 21, no. 1, pp. 324–337, Feb. [2013].
- [7] O. Gungor, J. Tan, C. E. Koksal, H. E. Gamal, and N. B. Shroff, "Joint power and secret key queue management for delay limited secure

- communication”, presented at the IEEE INFOCOM 2010, San Diego, CA, USA, Mar. [2010].
- [8] C. Capar, D. Goeckel, B. Liu, and D. Towsley, “Secret communication in large wireless networks without eavesdropper location information”, in Proc. IEEE INFOCOM, Orlando, FL, USA, Mar. [2012], pp. 1152–1160.
- [9] S. Sanghavi, D. Shah, and A. Willsky, “Message-passing for maximum weight independent set”, IEEE Trans. Inf. Theory, vol. 55, no. 11, pp. 4822–4834, Nov. [2009].

# LCF: A Temporal Approach to Link Prediction in Dynamic Social Networks

A. Jagadishwari. V  
Research Scholar, VTU,  
BMS College of Engineering, Bengaluru  
jiswariv@yahoo.com

B. Umadevi. V  
BMS College of Engineering,  
Bengaluru , umav.77@gmail.com

**Abstract—** Online Social Networks have become a prominent mode of communication and collaboration. Link Prediction is a major issue in Social Networks. Though ample methods are proposed to solve it, most of them take a static view of the network. Social Networks are dynamic in nature, this aspect has to be accounted. In this paper we propose a novel predictor LCF for Link Prediction in dynamic networks. In this method we view Social Networks as sequence of snapshots, each snapshot is the state of the network of a particular time period. Each edge of the network is assigned a weight based on its time stamp. We compute the LCF score for all node pairs in the network to predict the associations that may occur at a future time in the Social Network. We have also shown that our predictor outperforms the standard baseline methods for Link Prediction

## I. Introduction

The digital Social Media has brought about a revolution to Mankind. It has drastically changed the way people connect, communicate and collaborate. Today people meet, chat, discuss, debate and even do business through Social Media. Online Social Networks(OSN) such as Facebook, Twitter, Instagram, Flickr facilitate these interactions. The exponential growth of these OSN's has opened up new arenas of research. Social Network Analysis(SNA) is field of research which deals with the tools and strategies for the study of social networks. Link Prediction(LP) is one among the various problems that has been addressed by SNA. Link Prediction is a task of predicting future interactions that may occur in the social network. Fig. 1 Illustrates the problem, in the friendship network shown in the figure we need to predict if any of them who are not friends presently could connect with each other and become friends at a future time. For example in the friendship network of Fig 1 we might be interested in finding out if Alice and Bob or Sam and Jack could become friends in future.

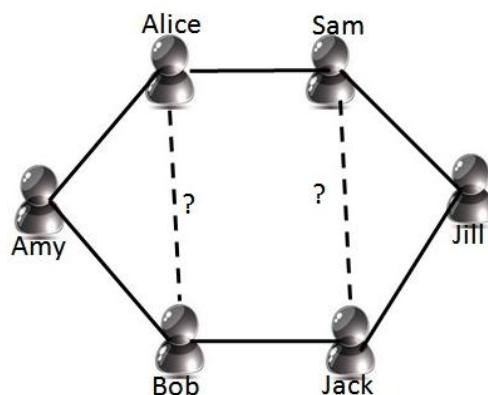


Fig 1. Illustration of Link Prediction Problem

Link Prediction can be applied in the study of network evolution [6], it also finds its application in recommendation system on social networks [3] to recommend friends on Facebook or Google+, to find employers or employees in professional networks such as LinkedIn, improving sales in E-commerce by suggesting products that customers may be interested to buy or even suggest some online shopping websites. Link prediction can also be employed in finding experts and collaborations in academic Social network [7] such as co-authorship networks. It can be used in biological networks such as protein to protein interaction and metabolic networks. It can be employed to unravel unknown connections in the terrorist networks [9]. The wide range of applicability of Link Prediction has generated a lot of interest among researchers and has attracted people from the fields of Computer science, Physics, Economics and Sociology [10]. This has served as a motivating factor for us to work on the problem. Though a lot of work has been done in this area, numerous methods have been proposed in the literature most of these methods take a static view of the Social Network. Social Networks are intrinsically dynamic and this nature has to be accounted. We have worked in this direction and come up with a novel method called the Latest Common Friend (LCF) predictor which embodies the temporal aspect of Social Networks. The contribution of our work is the following:

- A new predictor for Link Prediction in dynamic Social Networks LCF has been defined. We compute the score by using the time stamp of the edges and hence accommodating Time which is a key attribute of Social Networks.
- We define a Social Network to be an aggregate of a sequence of snapshots. The structure of each snapshot is decided based on the window size.
- The LCF predictor is compared with the standard baseline methods the Common Neighbor(CN) , Adamic Adar(AA) and Jaccard (JC) and shown that our method outperforms the baseline methods

## II. RELATED WORK

The methods that exist for Link Prediction can be classified as similarity based approaches, Path based approaches and Learning based approaches. Similarity based approaches uses the node's information or the topology of the network to predict links. libnen-Nowell [5] pioneered in proposing topology based prediction metrics that worked well on social networks. They worked on proximity metrics like Common Neighbor(CN), Adamic Adar(AA), Jaccard Coefficient(JC) etc and proved the prediction capability of these metrics. Some of the path based approaches are Local Path(LP) which uses the information of local paths between nodes to predict links and Katz measure which is based on ensemble of paths. Learning based approaches can further be classified as feature based or probabilistic models. In feature based approaches Link Prediction is treated as a typical binary classification problem and supervised classification learning models like Bayes or SVM's are used to solve it. These learning based approaches have difficulties in feature selection and are also computationally demanding as compared to similarity based approaches. Some probabilistic models such as Markov Random fields have been proposed. Most of these methods have been applied and tested on static networks. Since social networks are dynamic in nature , the relationships among the members of the network changes over time it becomes necessary for the prediction methods to imbibe the temporal aspect of these networks while predicting the future associations. Recently some approaches have been proposed to incorporate the temporal aspect of Social Networks while predicting the future links. In some of these methods the metrics used in the static networks are modified to suit dynamic networks. One such work is done by modifying the Common Neighbor metric, by finding the common neighbors within two hops [11] in the network. Similarly, weighted versions similarity metrics[2] have been employed on a time series graph. A random walk[1] based approach has been proposed for uncertain temporal networks in which similarity scores are computed for the node and a sub-graph surrounding the node. They integrate time and topological information to produce better results. Some Learning based approaches based on unsupervised feature learning method [8] has also been tested on Social Networks.

### III. PRELIMINARIES

#### A. Dynamic Social Networks

We are aware that with every passing minute a lot of activity happens on Social Media. A number of new profiles are been added, messages being sent, photos and videos been shared, this keeps changing the structure of Social Networks. The timing of the activities on the social networks is indispensable. Hence we define Dynamic Social Networks in the following way:

Let DSN be a Dynamic Social Network, with VD vertices and ED edges such that

$$V_D \in \{n_1, n_2, n_3, n_4, \dots, n_n\}$$

$$E_D \in \{e_1, e_2, e_3, e_4, \dots, e_n\}$$

For all  $e_i$  in  $E_D$  we have an time stamp attribute  $ts_{ei}$  indicating the time of creation of the edge  $e_i$ .

A Dynamic Social network is divided into N snapshots, each snapshot is represented as a graph. A graph of a particular snapshot will have all the vertices and edges contained in the network for a specific time period T. The number of snapshots depends on the window size (ws). If ws is 3 then the network will be divided into 3 snapshots with each snapshot having a time period T as shown below.

$$G_0 = (V_{t0}, E_{t0}) \text{ where } V_{t0} \subseteq V_D \text{ and } E_{t0} \subseteq E_D \text{ such that for all } e_i \text{ in } E_{t0} \text{ } ts_{ei} \text{ fall in the time period } t_0$$

$$G_1 = (V_{t1}, E_{t1}) \text{ where } V_{t1} \subseteq V_D \text{ and } E_{t1} \subseteq E_D \text{ such that for all } e_i \text{ in } E_{t1} \text{ } ts_{ei} \text{ fall in the time period } t_1$$

$$G_2 = (V_{t2}, E_{t2}) \text{ where } V_{t2} \subseteq V_D \text{ and } E_{t2} \subseteq E_D \text{ such that for all } e_i \text{ in } E_{t2} \text{ } ts_{ei} \text{ fall in the time period } t_2$$

$$G_{DSN} = G_0 \cup G_1 \cup G_2$$

here, the length of the time period  $t_0$ ,  $t_1$  and  $t_2$  are equal to T

#### B. Problem Definition

Given an undirected dynamic network  $G_D = \{V_D, E_D\}$  represented as a sequence of snapshots  $G_D = \{G_0, G_1, G_2, \dots, G_{Tg}\}$  where  $G_t = \{V_t, E_t\}$  and t ranging from 0 to T, for every node pair  $(x, y) \in V_t$  but  $\notin E_t$ , the link prediction task is to find if  $(x, y)$  will  $\in E_{t+1}$

#### IV. THE PROPOSED METHOD : LATEST COMMON FRIEND(LCF) PREDICTOR

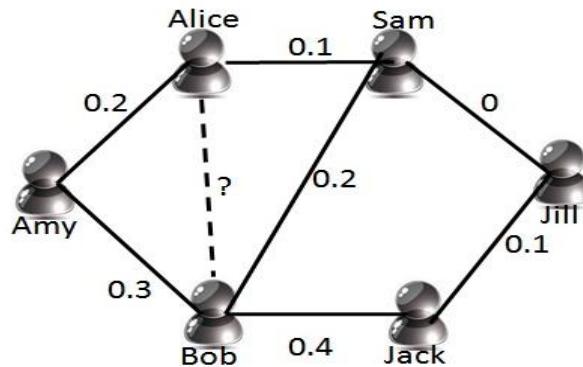


Fig 2. Illustration of LCF Score Computation

In the proposed method we define a new predictor the LCF, time stamp of the edges in the network is utilized in the computation of LCF score. Every edge in the network is assigned a weight based on the edge's time stamp. The edge with the oldest time stamp is assigned a weight 0 and the edge with the latest time stamp is assigned a weight 1. All the other edges are assigned weights between [0-1]. The LCF score of all node pairs in the network is computed by considering common friends of the node pair, such that the common friends who have a higher edge weight contribute significantly to the score. We compute the cumulative weight of common friends of a node pair as follow:

Let CF be the list of common friends of node pair (x , y) and N be the number of common friends,  $E_D$  is set of edges of a graph then:

$$Wt\_of\_CF = \sum_{K=0}^N Wt(x, CF[K]) + Wt(y, CF[K]) \quad (1)$$

$$LCF\_Score(x, y) = \| CF \| * Wt\_of\_CF(x, y) \quad (2)$$

$$Avg\_wt\_of\_Network = \frac{\sum_{M=0}^{\| E_D \|} Wt(E_D[M])}{\| E_D \|} \quad (3)$$

The LCF score of Alice and Bob is computed as follows:

Common Friends of Alice and Bob are ( Amy, Sam)

Weight of (Alice, Amy)= 0.3

Weight of (Bob, Amy) = 0.2

Weight of (Alice, Sam) = 0.2

Weight of (Bob, Sam) = 0.1

Cumulative weight of Friends of (Alice, Sam)=0.8

LCF score of (Alice, Sam)=  $2*0.8=1.6$

---

### Algorithm LCF\_Score

---

```

Input: Weighted Graphs { G1,G2,G3.....GN }
Output: LCF_Score of Gi
for every input graph { G1,G2,G3.....GN }
    E  $\leftarrow \sum$  wt of edges in Gi
    T  $\leftarrow$  Total Number of edges in Gi
    avg_wt  $\leftarrow$  E/T
    for each node pair( x , y ) in Gi
        CF(x,y)  $\leftarrow \Gamma(x) \cap \Gamma(y)$ 
        LCF  $\leftarrow$  0
        for each common friend k in CF(x,y)
            LCF  $\leftarrow$  LCF+1
            Cum_wt  $\leftarrow$  wt(x , k) + wt(y , k)
            if cum_wt > avg_wt
                LCF_Score(x,y)  $\leftarrow$  LCF * Cum_wt
            end if
        end for
    end for
    return LCF_Score
end for

```

---

The algorithm of the LCF predictor is outlined in Algorithm LCF\_Score. The input to the algorithm are weighted graphs, the weight to each edge of the graph is computed based on its time stamp. The LCF score for every pair of nodes of the graph is computed. The algorithm outputs the LCF score of every node pair of each input graph, based on the LCF score of edges in the Graph AUC values are computed.

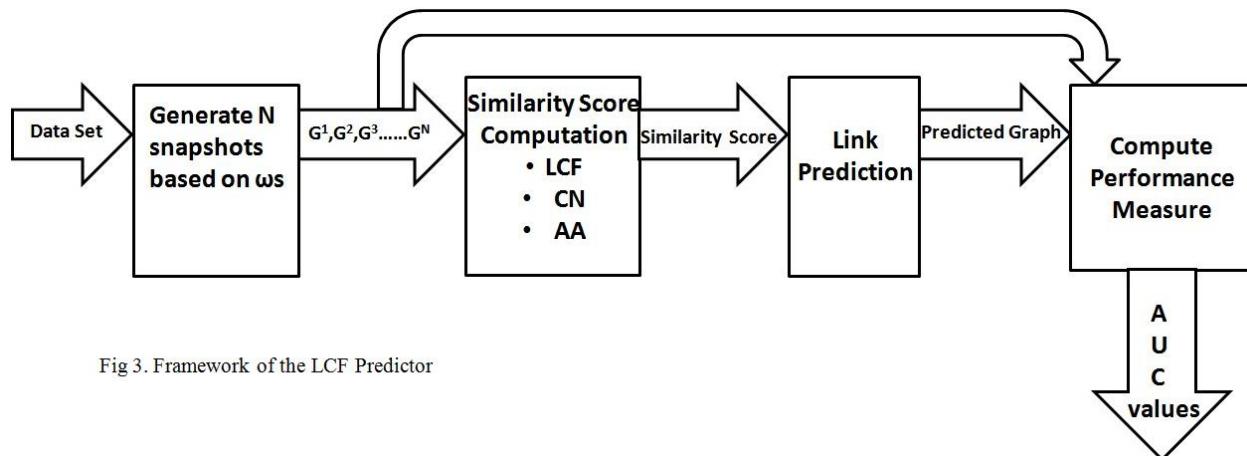


Fig 3. Framework of the LCF Predictor

## V. EMPIRICAL EVALUATION AND RESULTS

### A. Dynamic Social Networks

The Fig 3 shows the frame work of our Link Prediction method. There are three distinct phases:

1. Generate Network Snapshots: From every Data set generate N snapshots of the network with each snapshot of time period T, where value of T is decided based on the parameter ws.
2. Compute Similarity Scores: For every snapshot generated in phase 1, compute similarity scores based on LCF, CN, JC and AA.
3. Compute Performance Measure: Compare the scores of the edges computed in step 2 for a snapshot  $G^t$ , compare the scores of the existing edges and non existing edges in the snapshot  $G^{t+1}$  and compute AUC score

### B. Data Sets

The algorithm is tested on 8 different data-sets. All the data-sets are real time and temporal, they include the timestamps of every edge. There are multiple edges between two nodes one edge for each communication between the nodes. We consider two nodes to be connected if there is at least one communication either way. If there are multiple edges between two nodes the highest time stamp which indicate the latest time of communication is considered to be the time stamp of that edge. The Table 1 gives the description of the data-sets used to test our algorithms. The first four data-sets are communication data pertaining to the email communication among students of academic institutions or corporate communications. The other four are collaboration networks [4] pertaining to co-authorship or discussion forum through the internet on certain topics. The complete details of these data-sets can be found on the URLs provided in the Table 1.

Table 1. Data-set Description

Dataset	# of Nodes	# of Edges	Time Span (in months)	Source
College messages	1899	59835	193	<a href="http://snap.stanford.edu/data/CollegeMsg.html">http://snap.stanford.edu/data/CollegeMsg.html</a>
Enron	87,273	1,148,072	120	<a href="http://konect.uni-koblenz.de/networks/enron">http://konect.uni-koblenz.de/networks/enron</a>
Digg	30,398	87,627	1	<a href="http://konect.uni-koblenz.de/networks/munmun_digg_reply">http://konect.uni-koblenz.de/networks/munmun_digg_reply</a>
EU_ALL	265,214	420,045	803	<a href="http://konect.uni-koblenz.de/networks/email-EuAll">http://konect.uni-koblenz.de/networks/email-EuAll</a>
Math Overflow	24,818	506,550	78	<a href="https://snap.stanford.edu/data/sxmathoverflow.html">https://snap.stanford.edu/data/sxmathoverflow.html</a>
Ask Ubuntu	159,316	964,437	87	<a href="https://snap.stanford.edu/data/sx-askubuntu.html">https://snap.stanford.edu/data/sx-askubuntu.html</a>
Ca_HepPh	34,546	421,578	124	<a href="https://snap.stanford.edu/data/cit-HepPh.html">https://snap.stanford.edu/data/cit-HepPh.html</a>
Cit_HepTh	27,770	352,807	124	<a href="https://snap.stanford.edu/data/cit-HepTh.html">https://snap.stanford.edu/data/cit-HepTh.html</a>

### C. Results and Discussion

The proposed LCF algorithm was tested on 8 different datasets and the results were compared with the traditional methods Common Neighbor (CN), Jaccard (JC), and Adamic Adar (AA). The data-sets are divided into 10 snapshots. We compute the LCF score for a snapshot at time T0 ,T1,...T10. The LCF score is computed for all existing and non existing Links of all node pairs. The AUC value is computed using these scores. "AUC value can be interpreted as the probability that a randomly chosen existing link is given a higher score than a randomly chosen nonexistent link". Since we know all the edges that exists at time t1, we compare the scores of a randomly chosen existing link and a randomly chosen non existing link. The AUC value is computed as shown below:

Let n be the total number of comparisons made, if  $n'$  times an existent link had a higher score and  $n''$  times they have an equal score, then:

$$AUC = \frac{n' + 0.5 n''}{n'}$$

The Fig 4, shows the AUC values of 10 snapshots of the 4 communication data-sets, We have the our LCF algorithm performing better than the tradition ones except for one data set EU all, in which our CN performs on par with LCF. The Fig 5 shows the AUC values of the snapshots of collaboration networks, LCF performs very well in all four networks compared to the traditional methods. Finally in Fig 6, we compare the average AUC values for all 8 data-sets given by LCF and the other traditional methods, our LCF algorithm has higher average AUC values.

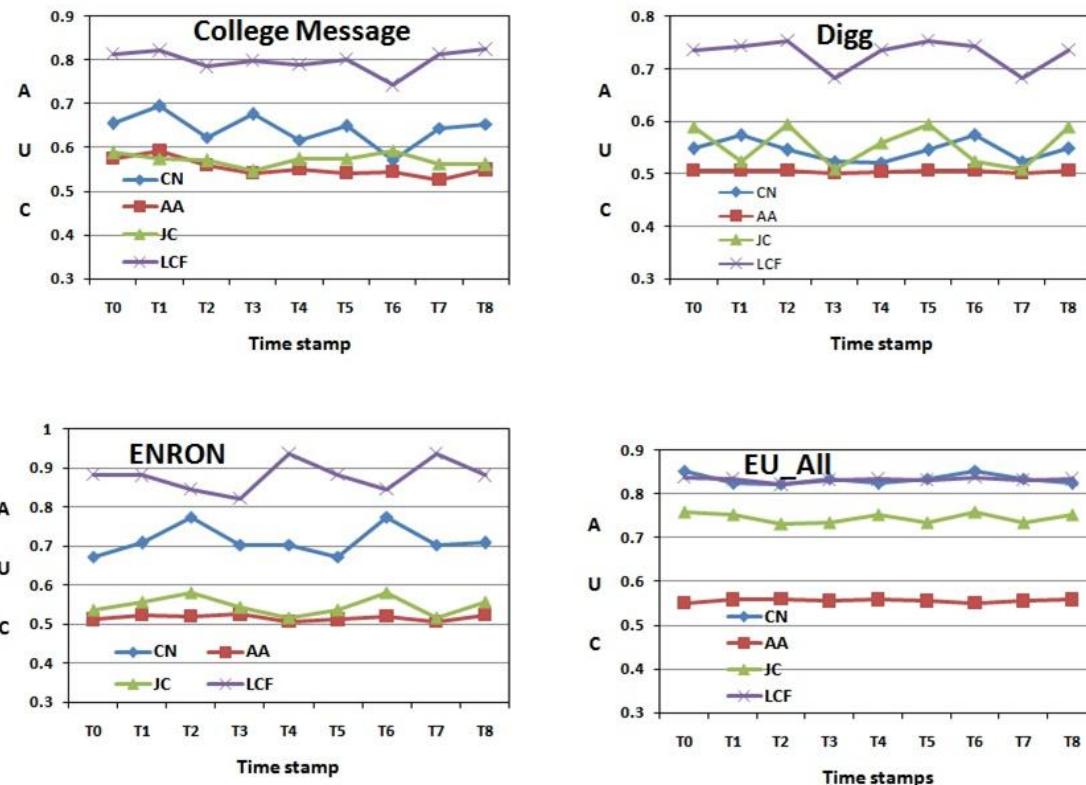


Fig 4. Performance of LCF and Traditional methods on Communication Networks

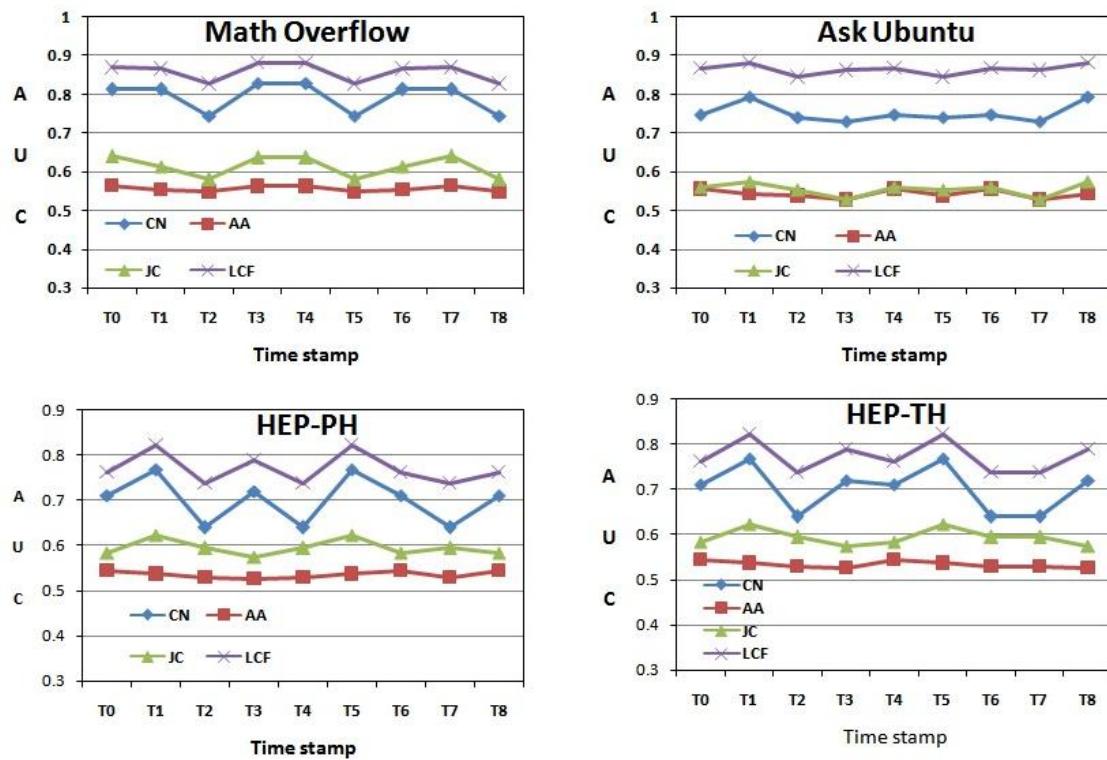


Fig 5. Performance of LCF and Traditional methods on Collaboration Networks

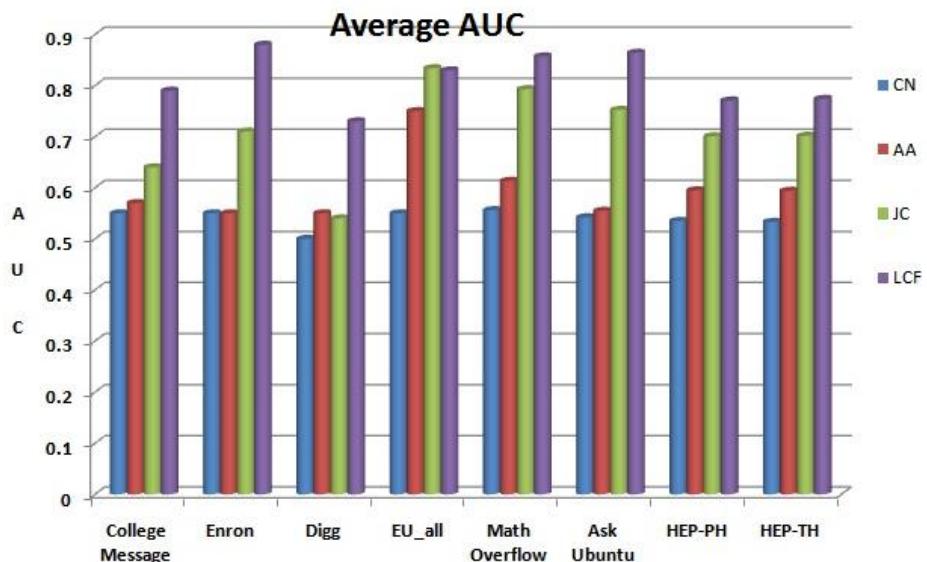


Fig 6. Average AUC Values of LCF and Traditional methods

## V. CONCLUSION

In this work the problem of Link Prediction in dynamic social networks was investigated. We have proposed a new algorithm for the problem. Time is a very important attribute in dynamic social networks hence in the proposed method we compute the scores for the edges by utilizing the time stamps of the edges. This algorithm has been tested on eight real world data-sets and we have shown that it performs better than the traditional Link Prediction algorithm.

## ACKNOWLEDGMENT

The work reported in this paper is supported by the college through the Technical Education Quality Improvement Programme [TEQIP-III] of the MHRD, Government of India.

## REFERENCES

- [1] Ahmed, N.M., Chen, L.: An efficient algorithm for link prediction in temporal uncertain social networks. *Information Sciences* 331, 120–136 (2016)
- [2] Ismail Güneş, Şule Gündüz-Öğüdücü, Zehra Çataltepe Link prediction using time series of neighborhood-based node similarity scores. *Data Mining and Knowledge Discovery* 30(1), 147–180 (2016)
- [3] Kuo, T.T., Yan, R., Huang, Y.Y., Kung, P.H., Lin, S.D.: Unsupervised link prediction using aggregative statistics on heterogeneous social networks. In: *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 775–783. ACM (2013)
- [4] Leskovec, J., Krevl, A.: SNAP Datasets: Stanford large network dataset collection. <http://snap.stanford.edu/data> (2014)
- [5] Liben-Nowell, D., Kleinberg, J.: The link-prediction problem for social networks. *Journal of the American society for information science and technology* 58(7), 1019–1031 (2007)
- [6] Liu, L., Zhou, T.: Link prediction in complex networks: A survey. *Physica A: Statistical Mechanics and its Applications* 390(6), 1150–1170 (2011)
- [7] Pavlov, M., Ichise, R.: Finding experts by link prediction in co-authorship networks. *Finding Experts on the Web with Semantics* 290, 42–55 (2007)
- [8] Rahman, M., Al Hasan, M.: Link prediction in dynamic networks using graphlet. In: *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 394–409. Springer (2016)
- [9] Ressler, S.: Social network analysis as an approach to combat terrorism: Past, present, and future research. *Homeland Security Affairs* 2(2), 1–10 (2006)
- [10] Wang, P., Xu, B., Wu, Y., Zhou, X.: Link prediction in social networks: the state-of-the-art. *Science China Information Sciences* 58(1), 1–38 (2015)
- [11] Yao, L., Wang, L., Pan, L., Yao, K.: Link prediction based on common-neighbors for dynamic social network. *Procedia Computer Science* 83, 82–89 (2016)

# An Implementation of a New Framework for Automatic Generation of Ontology and RDF to Real Time Web and Journal Data

Sudeepthi Govathoti<sup>1</sup>, M.S. Prasasd Babu<sup>2</sup>

*Research Scholar, Andhra University, Visakhapatnam &Assistant Professor, Anurag Group of Institutions, Hyderabad,  
Professor, Department of CS&SE, Andhra University, Visakhapatnam,  
[sudeepthicse@cvsr.ac.in](mailto:sudeepthicse@cvsr.ac.in)<sup>1</sup>, [profmspbabu@gmail.com](mailto:profmspbabu@gmail.com)<sup>2</sup>*

**Abstract—** Information on the web is tremendously increasing in recent years with the faster rate. This massive or voluminous data has driven intricate problems for information retrieval and knowledge management. As the data resides in a web with several forms, the Knowledge management in the web is a challenging task. Here the novel 'Semantic Web' concept may be used for understanding the web contents by the machine to offer intelligent services in an efficient way with a meaningful knowledge representation. The data retrieval in the traditional web source is focused on 'page ranking' techniques, whereas in the semantic web the data retrieval processes are based on the 'concept based learning'. The proposed work is aimed at the development of a new framework for automatic generation of ontology and RDF to some real time Web data, extracted from multiple repositories by tracing their URI's and Text Documents. Improved inverted indexing technique is applied for ontology generation and turtle notation is used for RDF notation. A program is written for validating the extracted data from multiple repositories by removing unwanted data and considering only the document section of the web page.

**Index Terms—** Semantic Web, Resource description framework, Ontology, Improved inverted indexing technique, Knowledge management.

## I. INTRODUCTION

World Wide Web (WWW) is considered as a global information repository that identifies documents and other web resources by Uniform Resource Locators, interlinked by hypertext links. Search engines are used to retrieve the information from the web. Data overburden is the most concerning issue in these days for the existing system. Evolution of web includes the web versions of web 1.0, 2.0 etc. In this series, the web version 3.0 is referred to as semantic web [1] is evolved as a knowledge management support across the globe. Search engines should be enriched with semantic web capabilities that analyze webpage content and provide more relevant results corresponding to the user query. Semantic web standards include resource description framework (RDF), web ontology, RDF Schema and rule

interchange format (RIF) for handling data. Resource description framework (RDF) provides a conceptual description of information for representing the web resources like Turtle syntax, N-Triples etc. Resource Description Framework (RDF) describes data on the Web in graph form [2]. Ontologies consist of the finite set of terms, relationships, constraints and axioms [3]. Ontologies have proven to be useful for effective knowledge modeling and information retrieval. The remaining paper is arranged as follows: In Section 2 the related work is presented. The proposed work and its methodology are discussed in Section 3 & 4. The results are presented in Section 5. Conclusions are given in Section 6.

## II. RELATED WORK

M.S.P.Babu et.al [4] provided the overview of some of the semantic search engines that yield unique search experience for users. Wilkinson et.al [5] proposed an information retrieval system using document structure. Amel Grissa Touzi et.al [6] suggested the Fuzzy Ontology of Data mining (FODM) for processing automated generation of ontologies in the domain of data mining. Amira Aloui et.al [7] implemented a plugin named "FO-FQ Tab plug-in", which can be integrated with protégé editor for building the fuzzy ontologies from large databases. To overcome the drawbacks of the existing system for accessing the related science information, M.S.P.Babu et.al [8] proposed a new framework for automatic generation of ontology and RDF for real-time web data. Tahani Alsabait et.al [9] developed the e-learning suite, with the set of questions designed using ontological representation. A.H.M.Rupasingha et.al [10] suggested that the performance of the ontology generation is always dependent on the specificity of the terms. Seongwook Youn et.al [11] discussed pros and cons of tools like protégé 2000, OilEd, Apollo, OntoLingua, Onto Edit, webODE, KAON, ICOM, DEO, webOnto that is used for ontology creation. Kgotatso Desmond Mogotlane et.al[12] presented a comparative study of plugins of protege tool like DB2OWL and Data Master.

### III. PROPOSED WORK

Semantic web capabilities like RDF & ontology are applied to enrich the knowledge. The proposed work is an implementation of the framework proposed by the authors [8]. The framework is designed with reference to the semantic web Stack. It is carried out in two phases, namely Data extraction phase and Data representation phase. Web scraping is performed using HTML parsing technique in data extraction phase by giving sample search query as an input to multiple repositories. DOM parsing and HTML parsing techniques are applied to validate the data retrieved from multiple repositories by considering only the document section of the webpage. Extensible markup language (XML) is the base for the semantic web representations; the validated information is converted into semi-structured notations by using XSD declaration from DOM tree and passed as an input for the next layers of the proposed framework. XML notation is given as an input to data representation phase. RDF notation is generated and represented in graphical form using Graphviz tool. A textual representation of RDF graph is provided using Turtle, the Terse RDF Triple Language. Improved Inverted Indexing technique is applied for ontological representation of words by excluding the stop words.

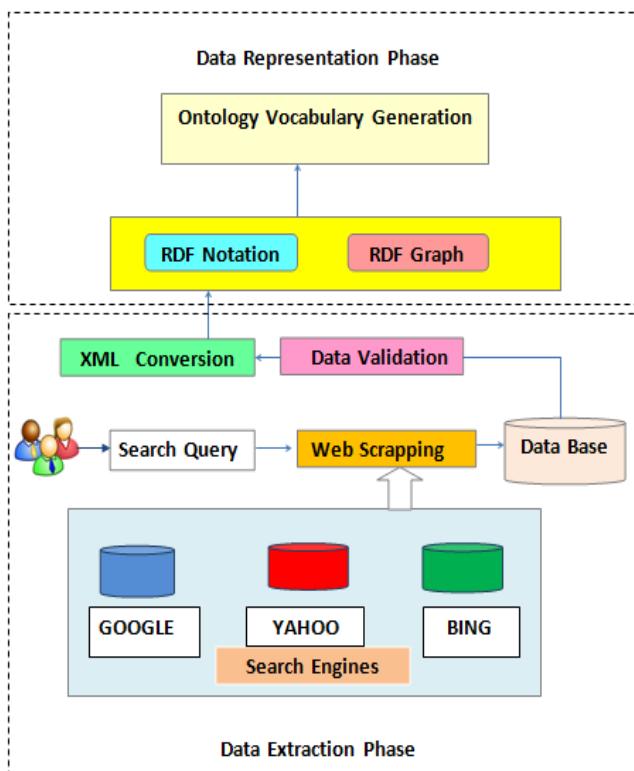


Figure 1: Proposed framework

### IV METHODOLOGY

Implementation of the framework proposed in Section III will be carried out in two phases namely data extraction and data representation phases. The details are given below.

#### Phase 1: Data extraction

Data extraction phase performs web scraping from multiple repositories and stores the scraped data into the database. The Data extraction phase is sub-divided into three steps namely web scraping, data validation, XML Conversion. The scraped data is further validated by removing the unwanted data in the considering document section of web page. The data stored in table format in the database, after the data validation process, is converted into the Semi-Structured Notations (i.e. XML Notations) and passed as an input to the data representation phase.

#### Step 1: Web Scraping

Web scraping, also referred as screen scraping or Web harvesting, is used to fetch and extract the data from a web document using HTML parsing techniques. Here Web pages are crawled and the content of the Web page is extracted. The data in the Web page includes three sections namely: Web page statistics bar, document section and descriptive section. The three items are stored in a database as three different attributes in a database table. HTML parsing technique is used for scraping data from the web documents is shown in Fig 2:

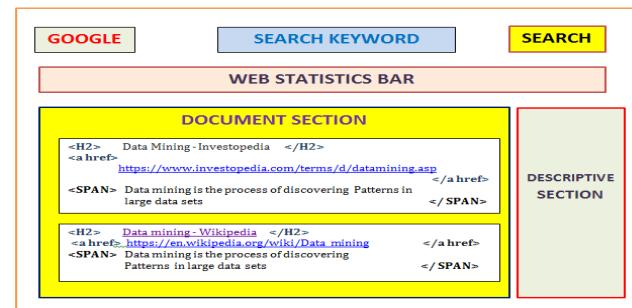


Figure 2: Content of the Web page

#### Step 2: Data Validation

In the Data Validation Step, the data collected from step 1 is validated using HTML and DOM parsing techniques. Here unwanted data is removed and the necessary portion of URLs is retained. In this step web status bar and descriptive sections are removed in the database table. The validated data is stored in a database. Document section displays the results in the form of page title, URL, Snippet (description) for the given search query.

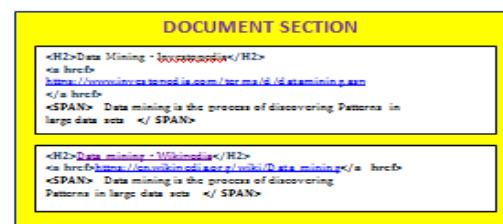


Figure 3: Content of the Web page after Data validation

Descriptive section provides the Wikipedia information about the input query. The Data validation process is carried out by considering only the document section of the web page as shown in Fig: 3.

### Step.3: XML Conversion

In XML Conversion Step the data, validated in step 2, is converted into a DOM tree using XML Schema Definition (XSD). The conversion is performed on data that is validated and stored in database by considering each individual field/ attribute into namespace convention. The XSD declaration of DOM tree has hierarchical structures which have root node, representing the search key word and three child nodes, representing Title, URL and Description respectively. The XSD declaration of the DOM tree with an example is shown in Fig 4:

```
<?xml version="1.0" encoding="UTF-8"?>
<bookstore>
  <book category="cooking">
    <title lang="en">Indian Cooking</title>
    <author>Sanjeev Kapoor</author>
    <year>2012</year>
    <price>80.00</price>
  </book>
  <book category="children">
    <title lang="en">Fair Tales</title>
    <author>K. Rowling</author>
    <year>2010</year>
    <price>50.99</price>
  </book>
</bookstore>
```

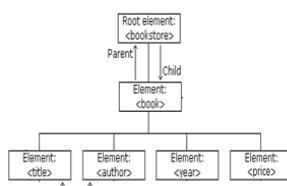


Figure 4: XSD Declaration of DOM tree

### Phase2: Data representation

Data extracted from steps 1,2 and 3 is maintained in an XML format and is given as an input to data representation phase. In Semantic Web architecture, the major source of data representation imposes RDF-ization and Ontology generation. Hence the data representation phase is sub- divided into two steps namely RDF-ization and ontology generation, which are explained in detail in step 4 & 5 respectively.

### Step 4: RDF-ization

The Resource Description Framework (RDF) is the basic building block in semantic web, promoting conceptual modeling of web data [13]. The RDF-ization process is carried out using Turtle notation and Graphviz tool. In this step the XML notation data stored in extraction phase is given as an input to RDF-ization. The RDF notation is visualized in the form of RDF graph using Graphviz tool. Decomposition of tuple creates a new blank node corresponding to the row and a new triple set is obtained. Each tuple in a relational database is decomposed as RDF triples, namely: the title is taken as subject, URL is considered as predicate and description is taken as object. A node can be a URI reference, literal or the blank node. The graph in Fig: 5 is an example of RDF-ization process of a semantic net.



Figure 5: RDF Triple

The triple is represented as a <subject, predicate, object> format by exploring the relationship among the nodes [14]. The XML conversion carried out in step 3 is represented in RDF syntax using Turtle notation from the convention specified in "<http://www.w3.org/1999/02/22-rdf-syntax-ns#>" as shown in Fig 6:

```
<?xml version="1.0" encoding="UTF-8"?> <b>"XML Declaration"</b>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <!DOCTYPE RDF Root element>
  <?xml version="1.0" encoding="UTF-8"?> <b>"Description of Resource identified"</b>
  <?xml version="1.0" encoding="UTF-8"?> <b>"Resource Property"</b>
  <?xml version="1.0" encoding="UTF-8"?>
  <rdf:Description rdf:about="http://www.semanticweb.org/rbss/hi">
    <rbss:keyword>hi</rbss:keyword>
    <rbss:title>Hi | Define Hi at Dictionary.com</rbss:title>
    <rbss:url>http://www.google.com/search?q=hi</rbss:url>
    <rbss:description>Hawaii (state) 2. Hawaiian Islands. Word of greeting, An American English, originally to attract attention (15c.), probably a variant of Middle English hy, hey (late 15c.) also an exclamation to call attention.</rbss:description>
  </rdf:Description>
</rdf:RDF>
```

Fig-6: RDF-ization

The <rdf: Description> element provides the description of resource identified by <rdf: about> attribute. The tags <rbss: title>, <rbss: keyword>, <rbss: URL> are the properties of the resource identified. The RDF represented in turtle notation is visualized in a graphical format using Graphviz tool .It is open source software that is used for generating graphs.

### Step 5: Ontology Generation

Ontology is defined as a formal specification of conceptualization of the domain of Interest. In ontology generation step, the RDF notation obtained from step 4 is used to create a vocabulary of words using improved inverted indexing algorithm. Improved Inverted Indexing algorithm is employed on real time web data collected from multiple repositories and text documents. The words from the description tags are extracted by excluding the stop words and frequency count/Term frequency (TF) of each word is maintained. The illustration of improved inverted indexing algorithm is presented as follows:

#### Algorithm: Improved inverted indexing

**Input:** Database D= {T<sub>1</sub>, T<sub>2</sub>...T<sub>n</sub>}, Storage Database

**Output:** Attributes {A<sub>1</sub>, A<sub>2</sub>...A<sub>n</sub>}, where A<sub>i</sub>, for i=1,2...n are representing ontology vocabulary.

**Parameters:** Swrd<sub>k</sub>= Array of Stop Words

attsL<sub>q</sub>= Snippet attribute

Words<sub>k</sub>= Words stemmed from snippet attribute

attsL<sub>f</sub>= Word frequencies after stemming

attsL= ontology along with the frequencies count.

1. Swrd<sub>k</sub>={α};
2. for i=0;i<=i+δ,j≤ D do
3. attsL<sub>q</sub>=Query Coverage(D,i);
4. Words<sub>k</sub>=Words Separate(attsL<sub>q</sub>,Swrd<sub>k</sub>);
5. attsL<sub>f</sub>= Words Usage frequency(D,attsL<sub>q</sub>,Words<sub>k</sub>);

```

6. attsL=attsLqU attsLf
7. f=highest_freq(attsL)
8. if (f<freq(attsL)) then
9. sort(Wordsk,freq(attsL))
10. end if
11. end for
12. return (Wordsk,freq(attsL))

```

## V. RESULTS AND DISCUSSION

The Semantic web stack proposed by Tim Berners Lee [15] is implemented by using the frame work proposed by the authors in section III. It is implemented in PHP version: 5 (Open Source scripting language) and MySQL version: 5 (open-source relational database management system) environments. It is tested on an input with test dataset comprising of sample search keywords. Response time is the amount of time that elapses from the receipt of the query until the results are displayed to user. Response time can be measured on server side or client side as shown in Fig 7.

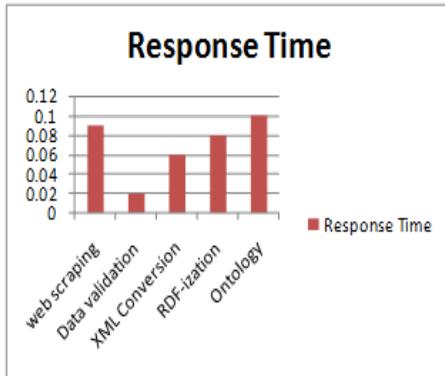


Figure 7: Response Time

Throughput is defined as number of queries executed per second (qps). Throughput and response time are observed for the set of retrieval operations with respect to the page load times. The performance of framework implemented with respect to throughput is shown in Fig 8.

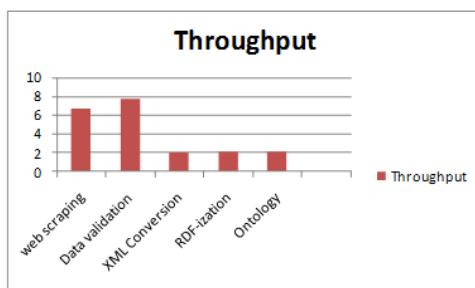


Figure 8: Throughput

A sample search query is given as an input and web scraping results are shown in Fig 9:

Figure 9: Web scraping results

Web scraping performance is evaluated by considering the following parameters like database size and count of URL's extracted which is shown in Fig 10:

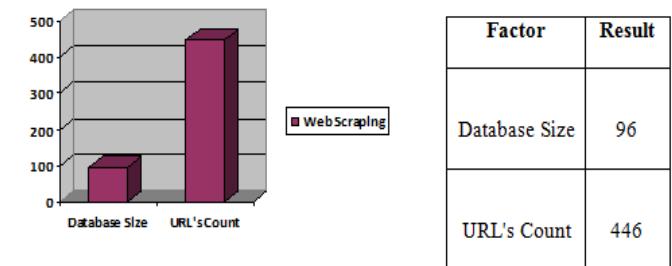


Figure 10: Web scraping analysis

Scraped data from multiple repositories is given as an input to data validation step. The validated data is obtained as an output to data validation process by applying HTML and DOM parsing technique. Data validation considers only the document section of a web page. Data validation results are shown in Fig 11:

Figure 11: Data validation results

The performance of data validation processing with wanted and unwanted data from the scraped data by considering the following parameters like database size and count of URL's is shown in Fig 12:

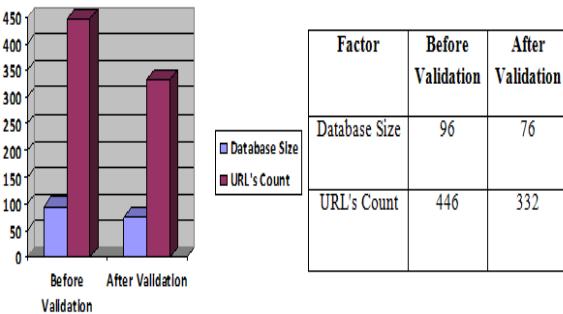


Figure12: Data Validation

The validated data stored in database is converted into the XML notations by applying XSD declaration as shown in Fig 13:

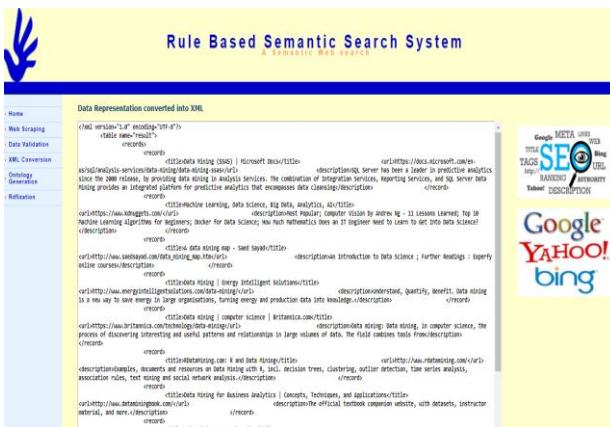


Figure13: XML Conversion

Resource Description Framework (RDF) is a recommended standard of World Wide Web Consortium (W3C) [16]. RDF representation of data in turtle form is shown in Fig 14:



Figure14: RDF-ization results Turtle form

RDF generation for sample relation named "testrdf" which has an attributes as <name, description, freq> is considered. The "testrdf" represents the relation name is considered as a class in an RDF graph and has set of three nodes that are connecting the testrdf in depth wise manner represents the tuple of a relation as shown in Fig 15:

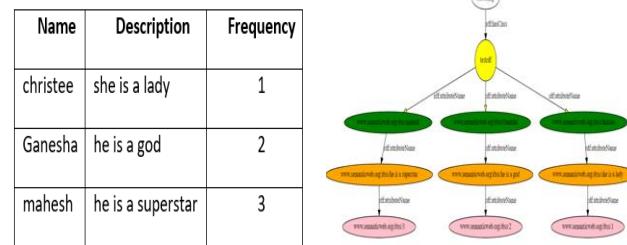


Figure15: RDF Graph generation

Ontology generation for the data obtained from multiple repositories as well as the text file. Improved inverted indexing technique is applied for extracting the words with their frequencies discarding the stop words, in the order of highest precedence. The result of ontology generation for real time web data with frequency is shown in Fig 16:

Title	URL	Description	ID	Word	Frequency
IEEE 2017	http://www.wieefina.lyearp	Web mining is the integration of information gather	1	web	28
Web Minin g	http://slog ix.in/proje cts	S-Logix offers Best Projects in Web Mining, final y	2	mining	22
Web Conte nta	https://w ww.cs.uic.edu/~l	Web content mining, Structured data extraction, se	3	data	16
What is We	https://w ww.techopedia.co	Web Mining Definition - Web mining in the process	4	information	12
What is We	https://w ww.scaleu nlimite	Web mining is the use of data mining techniques to	5	projects	12
Web Minin g	http://dm .rcs.unm.edu/Pap	Chapter 21 Web Mining @ Concepts, Applications,	6	mining,	11
ta Minin	http://ww w.tutorialspoint	Data Mining World Wide Web - Learn Data Min	7	techniques	8
			8	discover	6
			9	large	5
			10	patterns	5
			11	identify	5
			12	engineering,	4

Figure16: Ontology Generation for Real time web data

The highest frequency word is considered as a frequent search term for the purpose of rule framing using description logic. The rule mapping is done for the efficient retrieval operation which will be future work. The result of ontology generation for text document is shown in Fig 17.

Word	Frequency
Rule	1
Based	1
Semantic	1
Abstract	1
Now	1
days,	1
trending	1

Figure17: Ontology Generation for text document

## VI. CONCLUSION

The evolution of web has taken many forms namely web 1.0, web 2.0, web 3.0 , web 4.0 which lead to high-end information retrieval systems using semantic web. The existing traditional system collects the data from search engines is exhibiting average performance in retrieval. Implementation of

proposed framework for automatic generation of ontology's and RDF improves the performance of traditional search engines by incorporating semantic capabilities. It includes the application of HTML parsing technique, DOM parsing techniques and Turtle notation of graphviz tool. The algorithm improves information retrieval in Semantic Web and Expert Systems. The future work includes applying efficient cryptography for securing database and rule framing for the design of an expert system.

## REFERENCES

- [1] Sareh Aghaei, Mohammad Ali Nematbakhsh and Hadi Khosravi Farsani, "Evolution Of The World Wide Web: From Web 1.0 To Web 4.0", International Journal of Web & Semantic Technology (IJWesT) Vol.3, No.1, January 2012.
- [2] Abdeslem DENNAI, Sidi Mohammed BENSLIMANE,"Semantic Indexing of Web Documents Based on Domain Ontology", I.J. Information Technology and Computer Science, 2015, 02, 1-11.
- [3] Seema Redekar, Vishal Chekkala, Siddhappa Gouda, Swapnil Yalgude,"Web Search Engine Using Ontology Learning",International Journal of Innovative Research in Computer and Communication Engineering, Vol. 5, Issue 3, March 2017.
- [4] G Sudeepthi, G Anuradha, M Surendra Prasad Babu," A survey on semantic web search engine" International Journal of Computer Science Issues, 2012/3 IJCSI, Volume 9 Issue 2 Pages 241-245.
- [5] R. Wilkinson, "Effective retrieval of structured documents". (S.-V. New York, Ed.) Pages 311 – 317, 1994.
- [6] Amel Grissa Touzi, Hela Ben Massoud and Alaya Ayadi," Automatic Ontology Generation for Data Mining Using FCA and Clustering", arxiv.org, no. 1311.1764.
- [7] Amira Aloui, ENIT, Tunis, Tunisia,"A Fuzzy Ontology-Based Platform for Flexible Querying", International Journal of Service Science, Management, Engineering, and Technology, Vol.6, Issue 3, July-September 2015, pp 12-26.
- [8] Prof. M Surendra Prasad Babu, Sudeepthi Govathotu,"A Semantic Model for Building Integrated Ontology Databases", 7<sup>th</sup> IEEE International conference on software engineering and service science.
- [9] Tahani Alsubait, Bijan Parsia, Ulrike Sattler,"Ontology-Based Multiple Choice Question Generation", Kunstl Intell, Vol. 30, 2016, pp 183-188.
- [10] Rupasingha A. H. M. Rupasingha, "Improving Web Service Clustering through a Novel Ontology Generation Method by Domain Specificity ", IEEE 24th International Conference on Web Services, 2017.
- [11] Seongwook Youn, Anchit Arora, Preetham Chandrasekhar, Paavany Jayanty, Ashish Mistry and Shikha Sethi,"Survey about Ontology Development Tools for Ontology-based Knowledge Management".
- [12] Kgotsa Desmond Mogotsane, Jean Vincent Fonou-Dombeu,"Automatic Conversion of Relational Databases into Ontologies".
- [13] FatemeAbiri, Mohsen Kahani, FataneZarinkalam"An Entity Based RDF Indexing Schema Using Hadoop and HBase", 2011.
- [14] Faizan Shaikh, Usman A. Siddiqui, IramShahzadi, SyedUami, Zubair A. Shaik "SWISE: Semantic Web based Intelligent Search Engine", 2010.
- [15] Gopal Pandey," The Semantic Web: An Introduction and Issues", International Journal of Engineering Research and Applications, Vol. 2, Issue 1,Jan-Feb 2012, pp.780-786.
- [16] Resource Description Framework (RDF). Model and Syntax Specification. Technical report, W3C.



**Prof. M.S.Prasad Babu** was born on 12 - 08-1956 in Andhra Pradesh, India. He obtained his bachelors to doctoral degrees from Andhra University. He has 39 years of teaching and research experience. He guided 12 PhD's and 210 PG students for their thesis. He was the president of ICT section of ISCA in 2006-07. He attended about 50 National and International Conferences in

India and abroad and presented keynotes. He contributed about 250 papers National and International journals and conferences. He developed 10 international reputed Web portals. He won ISCA Young Scientist Award in 1986, State Best teacher award for engineering in 2015 and Dr. Sarvepalli Radhakrishnan Best Academician Award of Andhra University for 2014. He was the Conference Steering Chair for eight IEEE ICSESS of Beijing, China from 2010 to 2017. Prof Babu is presently working as Vice Principal of AU Engg College, Chairman, Faculty of Engineering of Andhra University and senior Professor in Computer Science & Systems Engineering discipline.



**G. SUDEEPTHI** was born in, Rajahmundry, Andhra Pradesh, India in 1983. She received B. TECH from National Institute of Technology Warangal, India in 2005 and M. Tech in Computer Science & Engineering from KLC Vijayawada, India in 2007. She is working as part time research scholar in the Dept. of CS & SE, Andhra University and as a Assistant Professor at Anurag Group of Institutions Hyderabad, India. She has got more than ten years of teaching experience. She Qualified FET-2011 Conducted by JNTUH. She contributed seven research papers in International journals and Presented one research paper at International conference in Beijing, China and another in Warangal, India.

# Nonnegative Garrote as a Variable Selection Method in Panel Data

Borislava Vrigazova  
Sofia University "St. Kliment Ohridski"

**Abstract:** In this research, we broaden the advantages of nonnegative garrote as a feature selection method and empirically show it provides comparable results to panel models. We compare nonnegative garrote to other variable selection methods like ridge, lasso and adaptive lasso and analyze their performance on a dataset, which we have previously analyzed in another research. We conclude by showing that the results from nonnegative garrote are comparable to the robustness checks applied to the panel models to validate statistically significant variables. We conclude that nonnegative garrote is a robust variable selection method for panel orthonormal data as it accounts for the fixed and random effects, which are present in panel datasets.

**Keywords:** nonnegative garrote, feature selection, panel models, fixed effects

## A. Introduction

Large number of variables and observations characterizes big datasets. They are ‘big’ in dimension and length. The challenge before researchers is to extract all valuable information from big data for statistical inference. Researchers can derive valuable information by identifying the significant variables and/or reducing the size of the sample. In data mining theory, these methods are known as dimensionality reduction.

Dimensionality reduction results in reducing the size of a dataset so that the researcher can extract all significant information for research. It includes two techniques – feature selection and feature extraction. Feature extraction transforms high-dimensional data into a space of fewer dimensions, which contain all significant characteristics of data for statistical modelling. The data transformation may be linear, as in principal component analysis (PCA), linear discriminant analysis (LDA) or nonlinear, as in non-linear LDA, kernel PCA. Feature selection, in contrast, reduces the number of variables by choosing those, which are statistically significant. As feature selection excludes variables from a statistical model, the reduction of dimensionality is performed by variable selection. What type of dimensionality reduction technique will be used depends on the data type, researcher’s goals and the quality of data.

In this research, we make comparison between the performance of variable selection methods like ridge, lasso, adaptive lasso and nonnegative garrote. Although these methods are popular in academic literature, we will test their performance on panel data to prove empirically that nonnegative garrote outperforms other feature selection methods in panel data. As the nonnegative garrote is the only variable selection method that accounts for fixed and random regressors, we show empirically that nonnegative garrote provides robust variable selection in datasets without multicollinearity. Thus, we broaden the set of advantages of nonnegative garrote. Next section reviews academic literature. Section 3 and 4 present the theoretical and empirical framework. Section 5 concludes.

## B. Literature overview

For many years, one of the most popular methods for variable selection has been the selection criteria of Mallows. During the 1960s, the Mallows’ C was new and modern tool to select significant variables in regression problems. As linear regression suggested a pool of  $2k$  possible models and only a few would profit good enough fit of data,

criterion for choice of the best model was necessary. Mallows [4] invented Mallows' Cp as criterion for rejecting unfit models to filter the ones, which provide best fit of data provided the best model had the smallest prediction error. Gorman and Toman [3] explored further the advantages of the selection criterion using visualization tools. The best fit between  $2^k$  models would select based on the combination of variables, which provided the smallest prediction error. The Mallows' criteria appeared to be simple for visual and numeric interpretation.

Due to computational exhaustion and other disadvantages of Mallows' criterion [5], researchers continued to look for other methods to select statistically significant variables in linear regression problems. The Russian mathematician Tikhonov, who worked on mathematical applications in physics, invented ridge regression as a solution to ill-posed problems. In 1962, Hoerl [2] applied ridge regression as a variable selection method in statistics for the first time. Unlike, Mallows' Cp criterion, ridge regression imposes penalty on regression coefficients. Some of the coefficients shrink to a number very close to zero, which filters out the variables with insignificant impact on the model. Hoerl and Kennard [6, 7, and 8] defined the ridge trace and its applications in orthogonal and nonorthogonal problems. Unsolved was the problem about a way of choosing the optimal value of the ridge penalty parameter. In 1979 generalized cross-validation (GCV) was proposed as a method for selecting the ridge parameter [9]. In modern research articles, ridge regression is applied with GCV to various statistical problems. Although it does not shrink regression coefficients to zero, it is used along with other variable selection methods for model validation. The use of ridge regression is broader which justifies its application nowadays in data analysis [10].

Ridge regression as a feature selection method is often compared to lasso regression [11]. Lasso was devised in 1996 [11] as a method which filters significant variables. Similar to ridge, lasso imposes penalty to regression coefficients but unlike ridge, shrinks some coefficients to zero. In 2006, Efron and Hastie [19] devised least angle regression to perform shrinkage. The advantages of lasso range from fast computation to being easily used to provide solution paths to other methods like lasso. Many software packages, for example, provide solutions for lasso regression based on lars paths [20].

Lasso [16] is based on another shrinkage method introduced a year earlier – nonnegative garrote [12]. Nonnegative garrote has an advantage to lasso as it can provide robust results in small and big datasets. The difference among ridge, lasso and nonnegative garrote appears to be in the type of constraint on parameters. Depending on the constraint, the methods yield different results. Although ridge and lasso use cross validation to select the value for the tuning parameter [9 and 13], the nonnegative garrote can use both cross validation and bootstrapping depending on the data type [12]. The two methods and their adaptations are widely used in various topics [13 and 14].

Zou [15] criticized the lasso approach for yielding a non-robust solution in big datasets. He described its disadvantages in details in [15] and proposed oracle method called adaptive lasso. Adaptive lasso overcomes some disadvantages of lasso by weighing the variables in a dataset. He theoretically justified his hypothesis that adaptive lasso performs more reliable variable selection than lasso. The properties of adaptive lasso allow variable selection in big datasets with heteroscedasticity [17] and for choice of generalized empirical likelihood estimators [19].

In 2005, Zou and Hastie introduced a hybrid between ridge and lasso called elastic net regularization [21]. The elastic net combines the penalty terms of lasso and ridge and outperforms lasso. The elastic net can be applied in

metric learning [22] and portfolio optimization [23]. Similar to lasso [24], elastic net can be reduced to support vector machine model for classification problems [25 and 26].

The evolution of variable selection methods has led to adaptations of these methods depending on the research goal and the quality of dataset. In the next section, we will provide the theoretical framework behind ridge, nonnegative garrote, lasso and adaptive lasso as they appear in the original research articles.

### C. Methodological Framework

In this section, we will present the theoretical framework behind the shrinkage methods. They all fulfill the task of minimizing a target given constraint on the coefficients. The shrinkage methods differ only by the type of the constraint.

Ridge regression [11]:

$$\widehat{\beta}_r = \arg \min \|y - \sum_{j=1}^p x_j \beta_j\|^2 + \lambda \sum_{j=1}^p \beta_j^2 \quad (1)$$

As  $\lambda$  goes towards zero, the ridge coefficients coincide with the coefficients of the OLS regression. When  $\lambda$  goes towards one, ridge coefficients shrink towards zero.

A special case of the estimator for an orthonormal matrix is

$$\widehat{\beta}_j^{\text{ridge}} = \frac{\widehat{\beta}_j^{\text{ols}}}{1 + \lambda} \quad (2)$$

Equation 2 shows that with the shrinkage of coefficients, variance is minimized but bias is introduced. Ridge regression rarely shrinks regression coefficients to zero. Instead, it shrinks OLS coefficients to values close to zero.

Lasso regression [11]:

$$\widehat{\beta}_l = \arg \min \|y - \sum_{j=1}^p x_j \beta_j\|^2 + \lambda \sum_{j=1}^p |\beta_j| \quad (3)$$

The lasso, on the other hand, shrinks some coefficients to zero, thus eliminating the statistically insignificant variables. The penalty  $\lambda \sum_{j=1}^p |\beta_j|$  is called L1 penalty. The penalty parameter  $\lambda$  determines the amount of shrinkage. When  $\lambda = 0$ , no shrinkage is performed and the lasso parameters equal the estimates of the OLS regression. With the increase of the value of  $\lambda$ , more parameters are excluded from the regression. When  $\lambda = \infty$ , theoretically all coefficients are removed.

A disadvantage of the least absolute selection operator (lasso) is its inability to perform robust variable selection in datasets, which are highly correlated. To solve this problem, Zou and Hastie [21] devised the elastic net, which is another regularization method, which combines penalty terms from ridge and lasso. The avid reader can review their article "Regularization and Variable Selection via the Elastic Net" for further details on elastic net.

Despite the advantage of lasso over ridge in terms of variables shrinkage, lasso is robust only in very big datasets. It fails to provide robust feature selection in smaller datasets. Adaptive lasso was devised [15] to overcome this disadvantage of lasso.

Adaptive lasso [15] :

$$\widehat{\beta}_{al} = \arg \min \|y - \sum_{j=1}^p x_j \beta_j\|^2 + \lambda \sum_{j=1}^p w_j |\beta_j| \quad (4)$$

The adaptive lasso estimator has similar structure as lasso but the difference is the each coefficients has weight in the penalty term. The tuning parameter behaves as  $\lambda$  in lasso. By weighing the coefficients, statisticians avoid spurious elimination of variables. Adaptive lasso can result in robust feature selection even in smaller datasets.

Choosing the tuning parameter in feature selection method is another important issue. We have used the k-fold cross validation as the underlying method for choosing the optimal value of the tuning parameter in ridge and lasso. The k-fold CV being the most widely used. The k-fold cross validation draws k different samples from a dataset and compares the models' error. The goal is to be chosen the value of the tuning parameter, which minimizes the CV error. The k-fold CV method is given by equation 5:

$$CV(\lambda) = \frac{1}{K} \sum_{k=1}^K E_k(\lambda) \quad (5),$$

Where K is the number of samples drawn from the dataset and k varies from one to K.

K-fold CV can be used for determination of the tuning parameter in the nonnegative garrote (eq. 6). However, nonnegative garrote uses k-fold CV when the assumption for the random independent variables is fulfilled. The nonnegative garrote can select significant variables in small and big datasets.

Nonnegative garrote [14, 29]

$$\arg \min \frac{1}{2} \|Y - Zd\|^2 + n\lambda \sum_{j=1}^p d_j \quad (6)$$

Where  $d_j > 0$  for all j and

$Z = (Z_1, \dots, Z_p)$ ,  $Z_j = X_j \hat{\beta}_j^{LS}$   
 $\hat{\beta}_j^{LS}$  being the least square estimate,  $\lambda$  is tuning parameter. The nonnegative garrote estimate of the regression coefficient is given by

$$\hat{\beta}_j^{LS}(\lambda) = d_j(\lambda) \hat{\beta}_j^{LS}, j = 1 \dots p \quad (7)$$

Under orthogonal designs the nonnegative garrote estimator can be expressed by

$$d_j(\lambda) = \left( 1 - \frac{\lambda}{\hat{\beta}_j^{LS}} \right), j = 1 \dots p \quad (8)$$

The shrinkage factor, as a result, will be close to one if the least square estimator is large. If the least square estimator is small, the shrinkage factor can reduce to zero.

The problem about the optimal value of the tuning parameter in the nonnegative garrote case can be solved either by k-fold cross validation or by little bootstrap procedure [1]. K – fold CV in nonnegative garrote, unlike lasso and ridge, is performed if the independent variables are assumed to random and uncorrelated. Lasso and ridge lack such an assumption. When nonnegative garrote assumes fixed independent variables, the little bootstrap procedure described in [4] can select the tuning parameter. When the X variables are random, the selection of the best tuning parameter can be performed by cross validation [1].

Nonnegative garrote differs from other variable selection methods like ridge and lasso in its ability to perform robust variable selection for various assumptions for X variables (random or fixed) [1]. This property of nonnegative garrote reminds of panel models where the panel OLS accounts for fixed or random effects. The equation of the panel models is:

$$Y_{it} = \beta f(x) + c_{it} + \varepsilon_{it} \quad (9),$$

Where  $c_{it}$  denotes the types of effects and  $\varepsilon_{it}$  marks the error term. Although academic literature has not investigated the connection between panel models and nonnegative garrote as a panel variable selection method, we believe there is such connection. More specifically, we have conducted a research of panel fixed effects model and compared the results of variable selection via nonnegative garrote with little bootstrap to conclude that nonnegative garrote for fixed X variables successfully performs feature selection in panels. We compared the results of the nonnegative garrote procedure to lasso, ridge and adaptive lasso and discovered that they fail in panel datasets.

Although we base our conclusion on one dataset and many experiments must be conducted, we believe this observation can be used to test the robustness of panel models and, on the other hand, it can be a time saving instrument for initial investigation of panel models.

Scientists apply panel models to datasets, which contain both time series and cross-sectional observations. The problem with panel datasets is the fact that the model should account for the effects of time and individual characteristics. To do that, scientists perform tests (Hausman [30], Pesaran [31]) to examine whether time and/or individual effects are present. They also investigate whether data are connected by factors that are characteristics of the dataset (fixed effects) or randomness is the underlying process in dataset (random effects) [28]. As a result, finding the right panel model is usually time-consuming process as it includes running a big amount of models and testing them for robustness.

We believe the similar results between panel models and the nonnegative garrote can be attributed to the similarities between panel fixed effects and little bootstrap in the X fixed case. In the X random case, the panel random effects can describe a procedure similar to cross validation. Breiman [1] describes the little bootstrap procedure as a method for choosing random samples with replacement from a dataset. The replacement corresponds with fixing the X variables to be a particular quantity [1]. The panel fixed effects models also assume X variables to be fixed quantities. Thus, the little bootstrap procedure generates random samples of fixed X and finds the tuning parameter of nonnegative garrote, which results in the smallest MSE in the fixed case. Then, the nonnegative garrote shrinks the coefficients of the fixed variables and selects only those, which are statistically significant. The fixed effects panel model, on the other hand, is based on panel OLS method, which estimates the coefficients of X variables with fixed quantities. The estimates of some coefficients can be statistically significant that leads to their elimination. Despite the fact that nonnegative garrote with bootstrap and panel fixed effects models choose the significant variables in different ways, the two of them address the problem of fixed quantities in X variables and should result in similar statistical significance.

Similar parallel between random panel effects models and nonnegative garrote with cross validation can be done. Random effects assume that X quantities are random. The panel model estimates the coefficients, based on a panel OLS method with random independent variables. Some of the coefficients become statistically significant. Cross validation in the nonnegative garrote, procedure chooses random samples from data, which, unlike the little bootstrap, are independent. In this way, cross validation accounts for randomness in data. Once cross validation estimates the MSE of many random independent samples, the tuning parameter for the nonnegative garrote can be selected and shrinkage can be performed. The statistically insignificant variables are shrunk to zero.

In the next section, we show our experiment on a panel dataset about property rights which contains fixed effects. Although our research needs to be extended with bigger amount of datasets, we believe that our results contribute to the practical advantages of the nonnegative garrote.

#### D. Results

We have carried out our experiment on a two datasets - property rights dataset, described in table 1 and a scoreboard of indicators for financial crisis in the EU described in [30].

We have analyzed dataset 1 in a research for determinants of property rights in a panel of data [29]. Our analysis shows that panel fixed effects models are present in the dataset. We have outlined the main determinants of property rights (model 1) and tested whether shadow economy (model 4), income (model 3) and gender inequality (model 2) affect the index of property rights. Using the panel OLS method, we have statistical significance of variables presented in table 1:

Table 1: OLS Panel method results:

	<b>Model 1</b>	<b>Model 2</b>	<b>Model 3</b>	<b>Model 4</b>
dbirth	-0.06** (-28.66)	-0.06** (-28.67)	-0.06** (-27.47)	-0.05* (-25.22)
drate	-0.10** (-26.17)	-0.10** (-26.16)	-0.10** (-25.81)	-0.11** (-27.16)
mortality	-0.05*** (-70.89)	-0.05*** (-72.08)	-0.05*** (-72.58)	-0.05*** (-70.78)
unempl	-0.02** (-27.96)	-0.02** (-27.96)	-0.02** (-25.91)	-0.02** (-28.54)
urban	0.01*** (35.26)	0.01*** (34.75)	0.01*** (32.54)	0.01*** (35.84)
dmilitary	0.26* (24.76)	0.27* (24.78)	0.27* (25.06)	0.26* (24.08)
dgender	-0.00 (-0.05)			
lginini			-0.09 (-0.51)	
dshadow				0.07*** (40.37)
R <sup>2</sup>	0.67	0.67	0.67	0.68
F-stat	135.59***	115.93***	116.55***	118.45*** 2000 - 2014
Data set	2000 -2014 Annual	2000 -2014 Annual	2000 -2014 Annual	2000 - 2014 Annual
N	420	420	420	419

Source: authors' calculations

According to the results of our experiment, income and gender inequality are statistically insignificant. The panel models lack cross sectional dependencies and multicollinearity and robustness checks have been made [29].

We have also performed variable selection on the transformed dataset (table 1) via lasso, ridge, adaptive lasso and the nonnegative garrote for fixed X variables. Table two presents the results of feature selection methods and compares them with panel models (1-4):

Table 2: Feature selection methods

	Lasso	Ridge	Adaptive Lasso	NNG	Panel models significance
dbirth	-0,05	-0,05	0	0,00	**/*
drate	-0,11	-0,13	-0,03	-0,05	**
mortality	-0,04	-0,04	-0,04	-0,48	***
unempl	-0,02	-0,02	-0,02	-0,22	**
dshadow	0,04	0,04	0,01	0,01	***
dgender	0	0,01	0	0,00	
infl	0	0	0	0,00	
loginternet	-0,02	-0,01	0	0,00	
dexpect	-0,05	-0,07	0	0,00	
emissions	0,01	0,01	0,01	0,00	
dhealth	0	0	0	0,00	
urban	0,01	0,01	0,01	0,33	***
lgini	-0,07	-0,11	0	0,00	
dmilitary	0,2	0	0,13	0,08	*

Source: authors' calculations

When we compare the results of the feature selection methods with the results of the panel models we see that lasso identifies 11 non-zero variables against seven significant variables from the panel models. Zou and Hastie [15] criticize lasso estimation for providing non-robust estimates in relatively small datasets and highly correlated variables. Although the variables lack multicollinearity and autocorrelation, the number of observations in the dataset is relatively small, so lasso fails to perform robust shrinkage of all insignificant variables. Another reason why lasso fails in this dataset is the presence of panel fixed effects, which are absent in cross-sectional and time series datasets. Lasso does not account for panel effects, fixed or random. Lasso was designed with the purpose to perform shrinkage of coefficients to zero in large datasets, most often applied on cross sectional data [11]. As a result, fixed and random effects in panels were not included in the theoretical framework of lasso and lasso fails to provide robust shrinkage in panel data.

The disadvantages of lasso have encouraged Zou and Hastie [15] to propose a weighted version of lasso, which is robust in the presence of multicollinearity and smaller number of observations. They called it adaptive lasso. In table 2, we see that the adaptive lasso outlines seven variables with non-zero coefficients and their amount coincides with the amount of significant variables in panel models. However, a more detailed analysis shows that the non-zero variables from the adaptive lasso are different from the statistically significant variables in the panel models. According to the adaptive lasso, the birth rate is not significant and emissions of carbon dioxide are significant, which differs from panel models. Similar to lasso, adaptive lasso does not account for panel effects. Ridge regression [2] results in similar estimations like lasso with more nonzero coefficients as it does not shrink to zero. From all three methods ridge is the least robust as it does not perform variable selection in panel data and brings no valuable information about the dataset.

Bearing these results in mind, we wanted to examine whether nonnegative garrote with little bootstrap can provide competitive results to panel models. As Leo Breiman [1] has introduced the advantages of nonnegative garrote as a feature selection method, we raised the question of whether it can be robust in panel data. Breiman [1] has proposed nonnegative garrote as an alternative of best subset selection and ridge. Later, Tibshirani and Hastie [11] have developed lasso based on the nonnegative garrote of Breiman. According to Breiman, nonnegative garrote performs robust variable selection not only in bigger datasets like lasso, but also in smaller ones. However, it fails when the dataset has multicollinearity and outliers. As our dataset lacks outliers and multicollinearity and has fixed effects, we ran nonnegative garrote with little bootstrap on our dataset.

As the nonnegative garrote is the only variable selection method, which proposes two selection procedures for the tuning parameter based on the assumptions for the X variables, it is reasonable to test it on panel datasets. As table two shows, the nonnegative garrote selects the same variables as the panel models. Although the coefficients have different values in nonnegative garrote and panels, the same variables appear to be statistically significant. According to the nonnegative garrote, the birth rate has zero coefficient, so it is statistically insignificant but in models, 1-3 it is significant at the 1% significance level and in model 4 the significance level is 10%, which shows the effects of birth rate may be negligible. Thus, this result is comparable with the result of the nonnegative garrote.

As the fixed effects in panel data imply fixed X variables and the little bootstrap in nonnegative garrote uses sampling with fixed variables, nonnegative garrote can be successfully applied as a panel feature selection method. In this research, we have validated the results from the panel models with panel GMM and robust covariance matrix (table 3). After we have confirmed the results, we have run nonnegative garrote. The results show that nonnegative garrote with little bootstrap was another way to validate the results of our research. Not only the selected variables via nonnegative garrote coincided with those from the panel models, but also with the results of the robustness checks. Nonnegative garrote proved to be useful not only for smaller datasets in variable selection in cross sectional data but also in panel data. As a result, nonnegative garrote with little bootstrap can be applied to fixed effects panel models without multicollinearity and outliers. Nonnegative garrote with cross validation can be applied to random effects panel models.

Table 3: Robustness checks

	<b>Model 5</b>	<b>Model 6</b>	<b>Model 7</b>	<b>Model 8</b>
DBIRTH	-0.06** (-1.93)	-0.08*** (-2.57)	-0.06** (-1.92)	-0.05* (-1.60)
DRATE	-0.11*** (-2.35)	-0.10*** (-2.53)	-0.10*** (-2.35)	-0.11*** (-2.41)
MORTALITY	-0.05*** (-15.93)	-0.05*** (-14.30)	-0.05*** (-15.33)	-0.05*** (-15.78)
UNEMPL	-0.02*** (-7.08)	-0.02*** (-7.56)	-0.02*** (-6.57)	-0.02*** (-7.07)
URBAN	0.01*** (10.21)	0.01*** (9.95)	0.01*** (10.01)	0.01*** (10.32)
DMILITARY	0.27*** (3.96)	0.26*** (3.83)	0.27*** (4.01)	0.26*** (3.86)
DGENDER		0.01		

		(0.81)		
LGINI		-0.09		
		(-1.19)		
DSHADOW			0.08***	
			(2.39)	
R <sup>2</sup>	0.67	0.63	0.67	0.68
Instrument rank	22	9	23	23
J-stat	139.34***	105.40***	149.22***	137.48***

Source: authors' calculations

Similar conclusion can be made from table 4 where we have analyzed the determinants of the change in the GDP in the context of EU scoreboard indicators for financial crisis [29]. Table 4 presents the results from the panel models with period and cross section fixed effects. Table 5 makes comparison between the panel model and the feature selection methods, which we have applied to the property rights dataset.

Table 4: Panel two-way fixed effects model for scoreboard of indicators for financial crisis in the EU

Dependent Variable: GDPCHANGE

Variable	Coefficient	Std. Error	t-Statistic	Prob.	
c	1.808557	0.736731	2.454842	0.0144	***
dcpi	0.055932	0.020464	2.733148	0.0065	***
dfincon	-0.109693	0.040897	-2.682185	0.0075	***
dgfcf	0.307502	0.053179	5.782384	0.0000	***
dgdebt	-0.010726	0.010249	-1.046570	0.2958	
dnip	-0.012993	0.005526	-2.351405	0.0191	***
dpsdebt	-0.001152	0.004332	-0.265959	0.7904	
drd	0.058476	0.193896	0.301584	0.7631	
dreer	0.005063	0.016126	0.313974	0.7537	
dresidential	0.137797	0.145171	0.949207	0.3429	
lactivity	0.188133	0.111589	1.685951	0.0924	*
ldcredit	-0.056197	0.118423	-0.474543	0.6353	
privatecf	0.021183	0.012114	1.748612	0.0809	*
riskpovst	0.033869	0.019763	1.713716	0.0871	*
totalfsliab	0.027452	0.007956	3.450620	0.0006	***
ulc	-0.019610	0.003790	-5.174630	0.0000	***
youthunemp	-0.031945	0.013864	-2.304101	0.0216	***
R-squared	0.620469				
F-statistic	14.04933				
Prob(F-statistic)	0.000000				

Source: authors' calculations

Table 5: Comparison feature selection / panel model for scoreboard of indicators

	ridge	lasso	adaptive	nng	panel significance
dcpi	0,01671	0,02767		0,07399	0,12600 ***
dfincon	-0,07761	-0,05544		0,00000	-1,82711 ***

dgfcf	0,23082	0,44817	0,58651	3,18178	***
dggdebt	-0,01821	0,01809	0,00000	0,00000	
dnip	-0,00686	-0,00558	0,00000	0,00368	***
dpsdebt	0,00131	0,00000	0,00000	0,00000	
drd	0,01177	0,00000	0,00000	0,00000	
dreer	0,01476	0,00000	0,00000	0,04267	
dresidential	0,33380	0,28351	0,00000	0,00000	
lactivity	0,02823	0,00000	0,00000	0,00515	*
ldcredit	-0,11780	-0,07883	0,00000	0,00000	
privatecf	0,02555	0,03043	0,02798	1,06233	*
riskpovst	-0,01291	0,00000	0,00000	0,32034	*
totalfsliab	0,01771	0,03002	0,03608	1,28388	***
ulc	-0,00515	-0,00988	-0,01187	0,00000	***
youthunemp	-0,01350	-0,00493	0,00000	0,73674	***

Source: authors' calculations

The column for panel significance in table 5 shows which variables are statistically significant according to the panel two-way fixed effects model. When we compare ridge with the panel model, it becomes clear that ridge failed to capture statistical significance, as all coefficients are different from zero. Ridge is not a reliable method for feature selection in panel datasets with fixed effects.

Similarly, to the property rights dataset, lasso shrunk some variables to zero. However, lasso identified variables, which are not significant in the panel model. For instance, the differenced government debt and the domestic credit are not significant in the panel in contrast to lasso. Lasso returned the log of activity rate (lactivity) as insignificant when, in fact, it is significant. Similar is the case with the risk of poverty (riskpovst). The lasso regression failed to perform feature selection successfully.

Adaptive lasso in contrast to lasso and ridge resulted in feature selection similar to the panel model. Some discrepancies can be observed, however. The adaptive lasso selected the differenced final consumption (dfincon), the net international investment position (nip), the log of activity rate (lactivity), the risk of poverty and the youth unemployment as insignificant when, in fact, they are significant. The accuracy of the adaptive lasso in the dataset is questionable.

When we look at the nng column, we see that the nonnegative garrote outlined as significant variables almost the same variables as the two-way panel model. The only discrepancy is in the dreer variable. The nonnegative garrote considers the real effective exchange rate to have nonzero coefficient while the panel model excludes it as insignificant variable. According to the nonnegative garrote, if the rate of increase in the real exchange rate becomes faster, the GDP will increase by 0.04%, which may be a very small, almost unobservable increase. In contrast to the other feature selection methods, the nonnegative garrote resulted in identifying all significant variables in the panel.

Similarly to the property rights dataset, the nonnegative garrote performed accurate feature selection in the panel scoreboard indicators for financial crisis. Despite the fact that different indicators and dependencies characterize the two datasets, the common line between them is the presence of fixed effects. The presence of fixed effects in panel models is modelled by OLS with fixed effects. As the little bootstrap captures the fixed effects, the nonnegative garrote successfully performs variable selection in panel data.

### E. Concluding remarks

Breiman [1] introduced the advantages of nonnegative garrote as a variable selection method. He proved its computational advantage to subset selection methods and its ability to perform robust variable selection in small and big datasets, unlike lasso and ridge. We broaden this set of advantages by showing that nonnegative garrote, unlike other variable selection methods, accounts for fixed X variables by little bootstrap procedure in fixed effects panel data and for random panel effects by cross validation. Although we lack sufficient empirical evidence, we believe our finding has an important impact on how panel data can be preprocessed.

### ACKNOWLEDGMENT

We thank prof. Ivan Ivanov<sup>1</sup> for the valuable comments and recommendations.

### REFERENCES

- [1]. Breiman Leo, Better Subset Regression Using the Nonnegative Garrote, Technometrics, Vol. 37. No 4, Nov. 1995, 373 – 384
- [2]. HOERL, A. E. (1962). Application of ridge analysis to regression problems. Chemical Engineering Progress 58, 54-59.
- [3]. Gorman, J. W. and Toman, R. J. (1966). Selection of variables for fitting equations to data. Technometrics 8, 27-51
- [4]. Breiman Leo, The Little Bootstrap and Other Methods for Dimensionality Selection in Regression: X-fixed Prediction Error, Journal of American Statistical association, 87, 738 -754
- [5]. Ronchetti Elvezio and Staudte Robert G., A Robust Version of Mallows's Cp, Journal of the American Statistical Association Vol. 89, No. 426 (Jun., 1994), pp. 550-559, DOI: 10.2307/2290858
- [6]. HOERL, A. E. and KENNARD, R. W. (1968). On regression analysis and biased estimation. Technometrics 10, 422-423.
- [7]. HOERL, A. E. and KENNARD, R. W. (1970). Ridge Regression. Applications to nonorthogonal problems. Technometrics 12
- [8]. HOERL, A. E. and KENNARD, R. W. (1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems, Source: Technometrics, Vol. 12, No. 1 (Feb., 1970), pp. 55-67
- [9]. Golub G., Heath M., Wahba G., Generalized Cross Validation as a Method for choosing the ridge parameter, Technometrics, Vol. 21, No2, 1979 May
- [10]. Peter Exterkate, Patrick J.F.Groenen, Christiaan Heijc Dickvan Dijk, Nonlinear forecasting with many predictors using kernel ridge regression, International Journal of Forecasting Volume 32, Issue 3, July–September 2016, Pages 736-753, <https://doi.org/10.1016/j.ijforecast.2015.11.017>
- [11]. Tibshirani, Robert, 1996, Regression shrinkage and selection via the lasso: a retrospective, J. R. Statist. Soc. B (1996) 73, Part 3, pp. 267-268, <http://statweb.stanford.edu/~tibs/lasso/lasso.pdf>
- [12]. Leo, Better Subset Regression Using the Nonnegative Garrote, Technometrics, Vol. 37. No 4, Nov. 1995, 373 – 384
- [13]. Woncheo Janga, Johan Lim, Nicole A.Lazar, Ji Meng Loh, DonghyeonYud, Some properties of generalized fused lasso and its applications to high dimensional data, Journal of the Korean Statistical Society Volume 44, Issue 3, September 2015, Pages 352-365, <https://doi.org/10.1016/j.jkss.2014.10.002>
- [14]. I.GijbelsI.Vrinssen, Robust nonnegative garrote variable selection in linear regression, Computational Statistics & Data Analysis Volume 85, May 2015, Pages 1-22, <https://doi.org/10.1016/j.csda.2014.11.009>
- [15]. Zou H., The Adaptive Lasso and Its Oracle Properties, Journal of the American Statistical Association December 2006, Vol. 101, No. 476, Theory and Methods, DOI 10.1198/016214506000000735
- [16]. Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B. Methodological, 58(1):267–288
- [17]. Wagener, J. and Dette, H. (2013). The adaptive lasso in high-dimensional sparse heteroscedastic models. Mathematical Methods of Statistics, 22(2):137–154
- [18]. Mehmet Caner , Qingliang Fan, Hybrid generalized empirical likelihood estimators: Instrument selection with adaptive lasso, Journal of Econometrics Volume 187, Issue 1, July 2015, Pages 256-274, <https://doi.org/10.1016/j.jeconom.2015.01.007>

- [19]. Efron, Bradley; Hastie, Trevor; Johnstone, Iain; Tibshirani, Robert (2004). "Least Angle Regression" (PDF). *Annals of Statistics*. 32 (2): pp. 407–499. doi:10.1214/00905360400000067. MR 2060166.
- [20]. LARS in Python [http://scikit-learn.org/stable/modules/generated/sklearn.linear\\_model.Lars.html](http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.Lars.html)
- [21]. Zou, Hui; Hastie, Trevor (2005). "Regularization and Variable Selection via the Elastic Net". *Journal of the Royal Statistical Society, Series B*: 301–320.
- [22]. Liu, Meizhu; Vemuri, Baba. "A robust and efficient doubly regularized metric learning approach". Proceedings of the 12th European conference on Computer Vision. Part IV: 646–659.
- [23]. Shen, Weiwei; Wang, Jun; Ma, Shiqian (2014). "Doubly Regularized Portfolio with Risk Minimization" (PDF). Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence: 1286–1292.
- [24]. Jaggi, Martin (2014). Suykens, Johan; Signoretto, Marco; Argyriou, Andreas, eds. An Equivalence between the Lasso and Support Vector Machines. Chapman and Hall/CRC.
- [25]. Wang, Li; Zhu, Ji; Zou, Hui (2006). "The doubly regularized support vector machine" (PDF). *Statistica Sinica*. 16: 589–615.
- [26]. Ming Yuan, Yi Lin, On the non-negative garrote estimator, *J. R. Statist. Soc. B* (2007) 69, Part 2, pp. 143–161
- [27]. Pesaran MH, 2004, General Diagnostics Tests for Cross Sectional Dependence in Panel Models, *Discus*
- [28]. Greene William H., 2003, Econometric analysis, 7th edition, Chapter 11: Models for Panel Data, Pearsonson Paper No. 1240, [repec.iza.org/dp1240.pdf](http://repec.iza.org/dp1240.pdf)
- [29]. Vrigazova B., Property rights: Factors Contributing to the Generation of Income from Ownership of Property, *Innovativity in Modelling and Analytics Journal of Research*, vol. 2,2017, pp. 22-39
- [30].[https://ec.europa.eu/info/business-economy-euro/economic-and-fiscal-policy-coordination/eu-economic-governance-monitoring-prevention-correction/macroeconomic-imbalance-procedure/scoreboard\\_en](https://ec.europa.eu/info/business-economy-euro/economic-and-fiscal-policy-coordination/eu-economic-governance-monitoring-prevention-correction/macroeconomic-imbalance-procedure/scoreboard_en)

---

<sup>1</sup> [https://www.researchgate.net/profile/Ivan\\_Ivanov20](https://www.researchgate.net/profile/Ivan_Ivanov20)

# IMPLEMENTING A SOLUTION TO THE CLOUD VENDOR LOCK-IN USING STANDARDIZED API

K V V N Prasanth<sup>1</sup>, K V V Satyanarayana<sup>2</sup>, V.Akhila<sup>3</sup>, M.Sahithya<sup>4</sup>, A.Sandeep Reddy<sup>5</sup>

Department of Computer Science and Engineering  
Koneru Lakshmaiah Education Foundation  
Vaddeswaram- 522502, Guntur Dt., AP

[kothamasu.prasanth@gmail.com](mailto:kothamasu.prasanth@gmail.com)<sup>1</sup>, [kopparti@kluniversity.in](mailto:kopparti@kluniversity.in)<sup>2</sup>, [akhila.vishnumolakala@gmail.com](mailto:akhila.vishnumolakala@gmail.com)<sup>3</sup>  
[sahithyareddy91@gmail.com](mailto:sahithyareddy91@gmail.com)<sup>4</sup>, [sandeepasr6@gmail.com](mailto:sandeepasr6@gmail.com)<sup>5</sup>

**ABSTRACT-** Cloud vendor lock-in is one of the major problems in cloud computing where the customer is locked to a particular vendor so that it will be difficult to migrate from one cloud to the other. The problem is that once an app has been developed based on a particular cloud service provider's API that app is bound to that provider as a result of which migration from one cloud to the other becomes more complex because of changes in architectures of different cloud service vendors[2]. The problem can be solved by providing a standardized way of interacting with cloud service providers taking many factors into consideration and by isolating each individual module involved in the cloud service provider's API and bringing out the common things and uniting them together so that in future any CSP will have to obey that specific standards and build their APIs without the need of creating a new standard that makes migration from/to that CSP complex.

**Key Words:** - Vendor lock-in, API, Migration, CSP

## 1. INTRODUCTION

Cloud computing is one of the most rapidly growing technologies in the world of Internet because it provides on-demand services to both end users as well as developers by providing various services such as SaaS, PaaS, and IaaS. As the popularity of cloud computing is growing higher and higher the problems are also growing much less at the same pace. One of the major problems is vendor lock-in where the customer is locked in to a particular vendor so that he/she has to use only that particular CSP when he uses it for

the first time, since the costs of migration will be higher both in terms of time and money. This serious issue is tackled in many different ways. One proposal is to provide a unified API which solves the vendor lock-in at PaaS and IaaS level and another proposal is the usage of Containers which are pretty much the same as Virtual machines but are more lightweight. These containers solve the problem at the IaaS level by combining the entire necessary environment needed to run and deploy an application leading to a smoother migration to another CSP but still it nevertheless faced the problem of locking in to that container and was also unable to solve the problem at PaaS level.

## 2. LITERATURE SURVEY [4]

The standardized API is based on an already existing system which is called as the Meta cloud. The standardized API derives some architecture of the Meta cloud. But provides a more fine grained control over the API and standardized for building most cloud services APIs. Some of the terms are below:

**Cloud:** Cloud is network of computers which provide several services to the end user where in the services are distributed to run and share data over a network i.e. the services can utilize the computing power of various computers connected to the cloud on-demand so that there is no wastage of any resources. Most of the applications are now moving to the cloud because of scalability.

**IaaS:** This stands for Infrastructure as a Service which means the cloud vendor provides only the infrastructure i.e. Hardware needed to build and deploy applications. The customer needs to install his own software on the top of the vendor's infrastructure.

**PaaS:** This stands for Platform-As-A-Service which means that the vendor provides an environment or a development platform to build and deploy an application. For example, the vendor may provide Debian Linux with Java installed over it. So, if the customer wants to develop a Java based application, he/she need not bother about the head ache installing Java software or Operating system into the vendor's computer.

**SaaS:** This stands for Software-As-A-Service which means that the vendor provides software for the end user to access to. Generally, these are CRM software that runs on the cloud. Some of the examples are Google Search, Facebook, and Salesforce etc. All of these come under Software as a service. These are most widely used.

Migration from one cloud to the other provider in case of SaaS may seem complex, but generally it is dependent completely upon the vendor and how well they facilitate migration.

## Different methods of migration Techniques

Live Migration can be a method that migrates the whole OS of one bodily system to a special. The digital system are migrated spirited while now not disrupting the appliances trolling on that.

The benefits of digital device migration encompass conservation of physical server, load equalization the various bodily servers and failure tolerance simply in case of sudden failure. The numerous virtual machine migration techniques are as follows

### Fault Tolerant Migration Techniques

Fault tolerance permits the virtual machines to continue its task even any part of device fails. This approach migrates the digital system from one physical server to a extraordinary physical server based totally upon the prediction of the failure befall, fault tolerant migration technique is to boost the supply of physical server and avoids performance degradation of applications .

### Load Balancing Migration Techniques

The weight balancing (or) leveling migration approach ambitions to distribute load across the distribute load across the physical servers to beautify the measurability of physical servers in cloud surroundings. the weight leveling aids in minimizing the resource intake, implementation of fail-over, improving measurability, keeping off

bottlenecks and over provisioning of resources and so forth.

### Energy Efficient Migration Techniques

The strength consumption of data middle is basically supported using the servers and their cooling structures. The servers generally up to seventy percentage of their most electricity consumption even at their low usage degree. Therefore, there's a demand for migration techniques that conserves the energy of servers by most fulfilling aid usage.

### LIVE VM MIGRATION IN CLOUD [3]

Live migration is an especially powerful tool for cluster and cloud administrator. Associate administrator will migrate OS instances with application so the machine will be used for physical functions.

There main 2 major approaches: Post-Copy memory and Pre-Copy memory migration. In the Post-copy memory migration approach it first suspends the migrating Virtual Server at the supply facet then once copies bottom processor state to the target host and resumes the virtual machine, and begins winning memory pages over the network from the supply node. There are two sections in Pre-copy approach: Pre-copy phase and Stop-and-Copy phase. In heat up VM memory migration section, the hypervisor copies all the memory pages from supply to destination whereas the VM remains running on the supply. If some memory pages amendment throughout memory copy method dirty pages, they'll be re-copied till the speed of recopied pages isn't but page change of state rate. In Stop and copy section, the VM are stopped in supply and also the remaining dirty pages are traced to the destination and VM are resumed in destination.

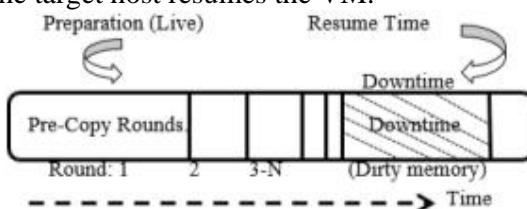
**Pre-Copy Phase [5]:** At this stage, the VM continues to run, whereas its memory is iteratively traced page wise from the supply to the target host. Iteratively means that, the algorithmic rule works in many rounds. It starts with transferring all active memory pages. As every spherical takes your time and within the in the meantime the VM remains running on the supply host, some pages could also be dirtied and need to be resent in a further spherical making certain memory consistency.

**Pre-Copy Termination Phase:** Without any stop condition, the iteratively pre-copy part could persevere indefinitely. Stop conditions rely

extremely on the look of the used hypervisor, however usually take one in all the subsequent thresholds into account: the amount of performed iterations exceeds a pre-defined threshold, the entire quantity of memory that has already been transmitted, exceeds a pre-defined threshold.

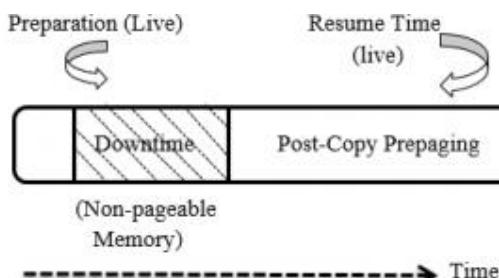
### Stop-and-Copy Phase

At this stage the hypervisor suspends the VM to prevent page dirtying and copies the remaining dirty pages and additionally because the state of the CPU registers to the destination host. Once the migration method is completed, the hypervisor on the target host resumes the VM.



**Post-Copy Approach:** Post-copy VM migration is initiated by suspending the VM at the supply, by suspending the VM a lowest subset of the execution state of the VM (CPU registers and non-pageable memory) is transferred to the target. The VM is then resumed at the target, albeit most of the memory state of the VM still resides at the source. At the target, once the VM tries to access pages that haven't nevertheless been transferred, it generates page-faults.

These two faults are cornered at the target and redirected towards the supply over particular network. Such faults are said as network faults. The supply host responds to the network-fault by causing the faulted page. Since every page fault of the running VM is redirected towards the supply, this system will degrade performance of applications running within the VM. However, pure demand-paging attended with techniques like pre-paging will scale back this impact by a good extent.



## Meta Cloud Architecture [1]

**Meta cloud:** This APIs can tackle the key-value pairs which stores the information about various cloud service providers which are referred by the developer while recoding. It is basically an architecture which contains Provisioning strategy, Knowledge base, Resource monitoring, Resource templates, Migration and deployment recipes, Meta cloud API and Meta cloud Proxy.

**Resource templates:** These describe cloud services required for running an application. This includes the services needed for an application, how each service interact with each other and any dependencies between each services, how well they are wired together etc.

**Migration and deployment recipes:** This includes what are the necessary tools or packages needed to migrate from one cloud to the other, any services required by the application.

**Resource monitoring:** This is a part of the API which provides monitoring various cloud services. The services can either go through the standardized REST API which makes use of the HTTP/HTTPS protocol or can come up with their own protocol. Monitoring includes how much bandwidth the application used, how many instances are deployed and running, active etc.

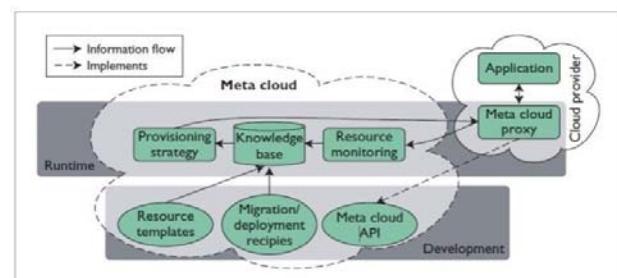


Figure 1. Conceptual meta cloud overview. Developers create cloud applications using meta cloud development components. The meta cloud runtime abstracts from provider specifics using proxy objects, and automates application life-cycle management.

## 3. PROPOSED SOLUTION FOR VENDOR LOCK-IN

Along with these above architecture specified in Winds of change, we have eliminated the need for Meta cloud proxy, Knowledge base. Since the focus of the paper is to provide standardized set of classes which can be used as the base for building any cloud service provider API. The API can well

be divided into two parts: one is the database part and the other is the Basic services part. We are trying to eliminate the lock-in of data to a particular CSP in the database part and in the basic part we are trying to eliminate the lock-in of the application to a specific vendor.

### 3.1 STANDARDS FOR BASIC CLOUD SERVICES

There is vast number of cloud services like Database-as-service, Java-As-Service etc. More or less, they can all be well grained into database and programming languages though there were many in SaaS architecture, the focus of the paper is to eliminate lock-in only in the PaaS model. Most of the databases are classified into either SQL Databases or NoSQL databases, each having their own model. Moreover, NOSQL databases can be classified into Key-Value stores, JSON models. When a PaaS service is chosen, migrating from it involves changing the connection code to that of the migrating CSP. This is not of a problem. The database architectural changes involve changing of the query commands. For example, SQL queries of Oracle differ with that of the MySQL in which case ORM frameworks can come into rescues which are already existent.

The one thing that definitely differs from one CSP to other CSP while migrating is the connection part of the code. This is the minimal change and can never be a problem of concern. Therefore, what needs to be standardized is the way of interacting with each CSP i.e. without changing the classes and much part of the code.

## 4. MODULES OF STANDARDIZED API

### 4.1 Database module

Here there will be a set of classes which are needed to interact with the database-as-service model of the Cloud service provider. To run an application database is very important and at times there may be a need to migrate to other database, may be to a different CSP.

In this case we are utilizing the existing ORM frameworks to interact with the databases. The ORM framework is one way of standardizing several databases since there is a conversion of a relation model to an object. Dependencies between table columns are specified in terms of dependencies between different instance variables

of different classes. SO for any database migration (only the relational databases), like from MySQL to Oracle or Oracle to SQL Server etc. can be easily facilitated by changing only the database specific connection code.

For NoSQL databases there are several formats. One is the Key-Value stores and the other is the JSON object notation format. Programming languages like Java already support the JSON object notation. So these set of classes can be used to convert Java objects to JSON format easily. The only part of work needed is that the cloud service providers for JSON type NOSQL databases must build their API depending upon these classes.

For Key-Value pair databases like Redis some classes in Java like Map and other utility classes can be served as a base. The part of the work again here is that the CSPs must conform to these standard classes. Since the architecture is already persistent, the work needed is to change the connection code which is discussed below.

### 4.2 Basic Services module

Here the basic services are actually more. Different CSPs provide different types of services and features. It is not possible to combine all of them. However, the API provides flexibility to the vendors to add additional features using these standardized APIs. For the basic connection codes, the standardized REST APIs are already present and are also supported by many CSPs and also many programming languages.

Apart from this, the CSPs can also choose their own protocols to interact with their services. In this case, the configuration details are enclosed in a separate file or database and are loaded from there. The API classes load those details and perform a request to those services as stated in the configuration. So, the basic services module eliminates the vendor lock-in problem of the application code.

### 4.3 Extensions module

The API also provides a way to facilitate adding additional capabilities so that the cloud service providers can extend the API and provide their own classes and interfaces to interact with additional features of their cloud service. The way the extensions module is provided is as a package of classes which facilitate the CSP to add additional classes. Any form of extension, could

be created as a package in the CSPs API and may involve several classes excluding the standardized API. Since any of the extension is some form of network activity (if it is an additional feature of the Cloud services itself, rather than the API) then, there are always protocols and classes necessary to access them.

For API specific additions, every extension class must inherit/implement the classes/interfaces (respectively) which term those classes in the API as extension classes. The extension module is nothing but a set of extension classes which are clubbed in a package. Any extension class that the CSP API would like to implement will have to implement the Extension interface.

## 5. EXPECTED RESULTS

The result of the standardized API would force the vendors to stick to specific standards but that does not restrict them from adding additional features to their cloud because of the extension module. This would result in an easier migration from one cloud to the other cloud. Since there is always information necessary to require migrating to a particular cloud and recipes are already stated in the architecture, the migration would only require changing the configuration file which deals with the information needed to connect to a particular cloud service provider. The standardized API made it simple to migrate to different cloud just like the JDBC did to migrate from one database to the other, just at the cost of changing the connection information.

## 6. CONCLUSION

The proposed solution for eliminate the Vendor lock-in solution must be implemented at the programming language level. Just like the Java programming language decided to come up with a JDBC API in the past to interact with databases from Java programming language which are conformed by all the database providers, this standardized API must also be implemented as a part of the programming language so that the CSPs conform to those and build their services based upon these set of classes. Rather than the customer locked in to the vendor, the vendor is locked in to this standardized API. The API covers most (if not all) of the cloud services by classifying them into basic categories and also provides a way to add additional functionality for

many other features that the Cloud service providers has to offer.

## 7. REFERENCES

- [1] Winds of Change – vendor lock-in to the meta cloud IEEE paper by FBenjaminSatzger, Waldemar Hummer, Christian Inzinger, Philipp Leitner, and SchahramDustdar • Vienna University of Technologyrom.
- [2] M. Armbrust et al., “A View of Cloud Computing,” *Comm. ACM*, vol. 53, no. 4,2010, pp. 50–58.
- [3] “Live VirtualMachine Migration Techniques in Cloud Computing,” [pdfs.semanticscholar.org/879e/626b7b8be0820c6b](http://pdfs.semanticscholar.org/879e/626b7b8be0820c6b)
- [4] Q. Zhang, L. Cheng, and R. Boutaba, “Cloud Computing: State-of-the-Art and Research Challenges,” *J. Internet Servicesand Applications*, vol. 1, no. 1, 2010, pp. 7–18.
- [5] A Survey of Virtual Machine Migration Techniques in Cloud Computing.

# Non-Separable Histogram Based Reversible Data Hiding Approach Using Inverse S-order and Skew Tent Map

Anoop Kumar Chaturvedi, Research Scholar  
Department of Computer Science & Engineering, University  
Institute of Technology  
Rajiv Gandhi Proudyogiki Vishwavidyalaya  
Bhopal, India  
[anoop.chaturvedi77@gmail.com](mailto:anoop.chaturvedi77@gmail.com)

Piyush Kumar Shukla, Assistant Professor  
Department of Computer Science & Engineering University  
Institute of Technology  
Rajiv Gandhi Proudyogiki Vishwavidyalaya  
Bhopal, India  
[pphdwss@gmail.com](mailto:pphdwss@gmail.com)

## Abstract

With the enhance in the digital media, modification and transfer of information is very easy. So this work focus on transferring data by hiding in the image. Here a robust approach is achieved by using the skew tent map as an encryption/ decryption algorithm at the sender and receiver side. In this work image is transformed into inverse S-order as the initial step of the work so little confusion can be created for the intruder. Here whole data hiding is done by modifying by using the modified histogram shifting method. This approach was utilized to the point that hiding information and image can be effectively recovered with no information loss. An investigation is done on the genuine dataset image. Assessment parameter esteems and demonstrates that the proposed work has kept up the SNR, PSNR, Throughput, Data Hiding Execution Time and

Extraction Time values with high security of the information.

*Keywords—Digital data hiding, Encryption, Histogram, Image Processing.*

## 1. Introduction

As the web is developing definitely clients are drawn in by different specialist organizations step by step. Some of online shops, computerized showcasing, informal organization [21, 23] and so on. This simple access prompts change the proprietorship effectively, as clients can steal other work and make computerized printed with their names. In any case, this innovation offer ascent to new issue of piracy [12, 18].

To conquer this issue numerous approaches were recommended and restrictive of the advanced information is protected. So to defeat this distinctive strategies are used for safeguarding the restrictive of

the proprietor. Out of many methodologies, advanced information inserting which is otherwise called computerized Data Hiding assumes a critical part. Keeping in mind the end goal to give proprietorship of the information proprietor, advanced information was implanted into the image, video, or information as in [11, 19, 23].

One of the real fundamentals of information hiding is that the hidden information must be hazy. The utilization of stenography has many points of interest and are extremely helpful in computerized picture handling which makes them appropriate for a wide collection of uses. In this cutting edge region, digital incredible comfort in transmitting a lot of information in various parts of the world [1, 15]. In any case, the wellbeing and security of long separation correspondence remain an issue. Keeping in mind the end goal to take care of this issue of security and wellbeing has promoted the advancement of stenography plans. Stenography is not quite the same as watermarking and cryptography [2, 16]. The fundamental target of stenography is to conceal the presence of the message itself, which makes it troublesome for a spectator to make sense of where precisely the message is. Then again, cryptographic systems have a tendency to secure correspondences by changing the information into a frame with the goal that it can't be comprehended by a meddler [6, 22]. Also, in watermarking logo is more critical than data. Stenography is the sort of concealed

correspondence that signifies "secured expressing" [5, 10] (from the Greek words stego or "secured" and graphos or "to write").

Information hiding is the procedure to cover information in a cover media. In this way, the information, concealing procedure contains two sorts of information [9, 13], embedded information and cover media information. The information is transmitted by implanting it inside Images, which enhances information security. The information, concealing strategy in which the reversibility can be accomplished is called Reversible information hiding [14, 17]. This method is used to improve the security of the cover Image [21] in encryption. Reversible image [21] data hiding (RIDH) [22] is one strategy for the information, concealing procedure, which ensures that the cover picture is recreated flawlessly after the removal of the implanted message. The reversibility of this technique makes the information, concealing methodology attractive in the basic situations, e.g., remote detecting, military, law, crime scene investigation, medical picture sharing [22, 24] and copyright confirmation, where the original cover picture is required after remaking [1].

## 2. Related Work

In [18] digital data were embedded in the selected portion of the image where the edge region was selected for embedding. Here paper has developed a new approach to finding pixel representing edges.

By utilizing Dam and BCV approach image was segmented into edge and non edge region. One drawback of this work that it was done on binary image only means data hiding can be done by this method for binary image only. This lead to one more limitation that is embedded data should of binary form only. With above issues image was highly robust against different types of attacks like filter [7, 16], noise, etc.

In [4] author has extended the work done in [18] by increasing the overall capacity of the embedding data space. Here at Dam and BCV technique author start looking at the surrounding of the edge region pixel. So the overall capacity of the data hiding was drastically increased in this paper. Here even after embedding more data embedded image was robust against different types of attack as well.

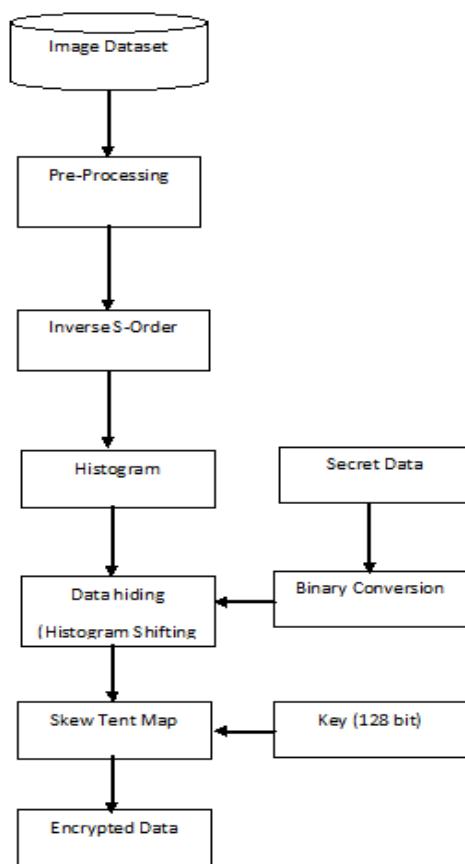
In [21] self embedding concept was proposed by the authors where the image itself generate the data for embedding while in order to protect data in network fountain codes were developed for lost packed regeneration. As in fountain codes more than one required packet format was sent on the network, which help in regenerating the missed or corrupt data packets. Here work has great limitation being that after embedding the image is not available in original format before extraction. So main purpose of this work is for transferring the data packet from sender to receiver only.

In [20] same concept of image Data Hiding self generation was done, here image was so utilized that it generates its own Data Hiding information. This paper center of attention on the image development where the spatial area was utilized for inserting the digital data as a carrier object. At the same time, similar information is required at the receiver which help in finding the digital data back. But to wrap both intra-code block and inter-code block method utilizes.

In [12] authors utilize the DWT feature for finding the pixel value for embedding. While in order to increase the randomness in the embedding the selection of images was not sequential but it would utilize the random Gaussian function for selecting pixel of different position. At the receiver side with the help of some supporting information it was found that Data Hiding [3, 8] was extracted from the image. Here it was obtained that both Data Hiding and image got reverse at the receiving end. In [25] author adopts KSVD technique for embedding the digital data. Here by utilizing the RC\$ algorithm encryption of the digital data was done. Here one dictionary was maintained at the receiver and the transmitter end for reducing the size of the carrier signal. In this work after embedding some vacant space between the data was utilized in the data embedding. This work has given freedom for the extraction of image or digital data or both in any order.

### 3. Proposed Methodology

Center of attention of this work was to hide digital data in the image. The whole work was done in two steps of embedding digital data and extraction of digital data. Here it is desired that while extracting digital data [21, 25]. In Fig. 1 whole embedding work block diagram is described.



**Fig.1 Block diagram of proposed work.**

#### 3.1 Pre-Processing

A collection of different number on fixed range represents a type of image format. So reading pixel values of that image matrix is done in this step of the proposed model.

As whole work focus on the image which has a pixel value in the scope of 0-255. So examine an image implies building a framework of the same. Measurements of the image at that point fill the matrix cell to the pixel value of the image at the cell in the grid.

#### 3.2 Inverse S-Order

In this step all the color channels Red, Green and Blue are into single dimension matrix or vector S. Here as per inverses s-order sequence of pixel value is inserted in the S vector. This can be understood by below example where fig. 2 (a) represent original image and fig. 2 (b) represent the inverse S-order of the matrix.

4	5	6	4
5	6	6	6
4	7	8	9
5	4	6	9

(a)

4	5	6	4
6	6	6	5
4	7	8	9
9	6	4	5

(b)

$$S = [4, 5, 6, 4, 6, 6, 5, 4, 7, 8, 9, 5, 4, 6, 9]$$

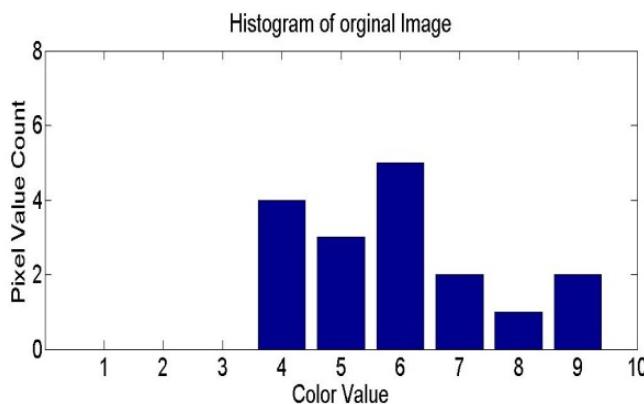
**Fig. 2 Inverse S-order representation.**

In this way whole image pixel values are arranged in the single vector S where the order of the pixel values is the inverse S-order. In case of the color

image, the first red matrix is insert than green is inserted and at last, blue matrix is inserted.

### 3.3 Image Histogram

In this step S vector obtained after inverse S-order is used where histogram of the image is found at one bins. This can be understood as let the scale of color in fig. 2 is 1 to 10, than count of each pixel value is done in the image. So as per above S vector  $H_i = [0, 0, 0, 4, 3, 5, 2, 1, 2, 0]$  where H represents the color pixel value count and i represent the position in the H matrix with color value.



**Fig. 3 Histogram of the original image.**

### 3.4 Histogram Shifting and Data Hiding

In order to make reversible data hiding this work adopt a histogram shifting method for data hiding in the image. From above step pixel value with number of presence is obtained where pixel having largest presence or the highest peak in the histogram is  $P = \{6\}$ . In similar fashion pixel having a zero presence in the image is  $Z = \{1, 2, 3, 10\}$ .

Histogram shifting is obtained by manipulating the peak value with zero presence pixel value, but this makes one limitation that numbers of data hiding bits are less. This can be understood as  $P = \{6\}$  where pixel value 6 is present in 5 locations of the S vector, so maximum 6 bit data can be hidden in this image carrier. So in order to increase the number of positions in the image proposed work has included other peak of the histogram for increasing the hiding capacity. This can be understood as if peak vector includes other pixel values let  $P = \{6, 4, 5, 7\}$  than total 12 bits can be hide in the image while replacement of the peak value are done by its zero value vector  $Z = \{1, 2, 3, 10\}$ .

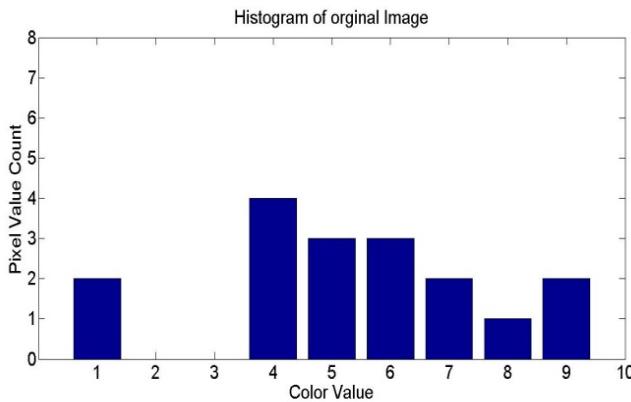
### 3.5 Data Hiding

Here histogram shifting is done in hiding each bit of the data. This shifting means replacing the peak pixel values with its corresponding zero pixel value. Let hiding data be  $H = [1, 0, 0, 1]$ . As per histogram shifting if bit 1 come in hiding data than the peak value to remain unaffected while when bit 0 come in hiding data than replace peak value with a zero value.

$$S = [4, 5, 6, 4, 6, 6, 5, 4, 7, 8, 9, 5, 4, 6, 9]$$

$$\begin{matrix} 1 & 0 & 0 & 1 \end{matrix}$$

$$HS = [4, 5, 6, 4, 1, 1, 6, 5, 4, 7, 8, 9, 5, 4, 6, 9]$$



**Fig. 4 Histogram of the original image.**

So the steps of hiding the data are

Input: S, H, P, Z

Output: HS // Hided Data

1. Pos=1
2. c=1 // c peak value position
3. Loop 1:H // For each hiding bit
4. p←P[c]
5. z←Z[c]
6. If H==0 then
7. Loop pos : n // n number of pixel values in S
8. If S[pos] = p
9. Jump step 12
10. End If
11. End Loop
12. S[pos] ←z
13. End If
14. Pos←pos+1
15. End Loop

### 3.6 Skew Tent Map

Finally obtained HS vector which contains secret data is handled in this step for encryption. Here Skew tent map algorithm takes 128 bit input as the key and HS vector as the plaintext. Before encryption some basic steps of skew tent need to do for getting various constants.

Step 1. Key 128 bits are divided into four sub keys termed as K1, K2, K3, K4 where each sub key contain 32 bits.

Step 2. To ascertain the underlying condition X0 for the principal skew tent guide, pick any two pieces of session keys i.e. K1 and K3. Now process a genuine number X01 utilizing the XOR operation between them:

$$X_{01} = \text{XOR}(K_1, K_3)$$

Further register another genuine number X02 as takes after Where n speak to a number of keys.

$$X_{02} = \frac{\sum_{i=1}^n k_i}{128}$$

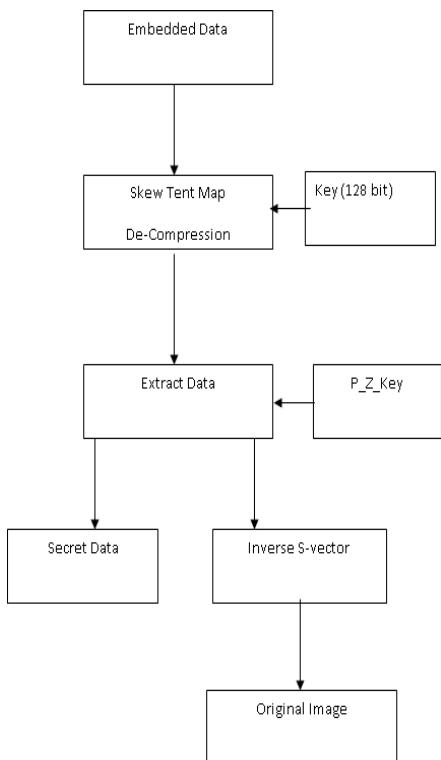
Where n represent a number of keys.

$$X_0 = (X_{01} + X_{02}) \bmod 128$$

With the underlying condition Xi by T times and refreshes Xi to the most recent status and also repeats the second skew tent map but one time for Yi given as below:

$$Y_i = T(X_i)$$

where  $T_i = T_0$  for  $i=1$  and  $T_i = T_{i-1}$



**Fig.5 Block outline of information extraction at the recipient end.**

Now the updated values of  $X_i$  and  $Y_i$  are used to encrypt and decrypt the  $i$ th plaintext and cipher text block given as below:

$$C_i = \text{XOR } (P_i, C_{i-1}, X_i) \text{ Encryption formula}$$

$$P_i = \text{XOR } (C_i, C_{i-1}, Y_i) \text{ Decryption}$$

### 3.7 Extraction steps

In this extraction steps receiver can extract data and image by using an above block diagram.

### 3.8 Extraction of Image

This section of proposed work is for image extraction at the receiver side. Here first skew tent map decryption algorithm is used for getting decrypted data. So resultant series obtained is taken as input in histogram shifting where peak and zero key pair is passing. After this hiding data is obtained.

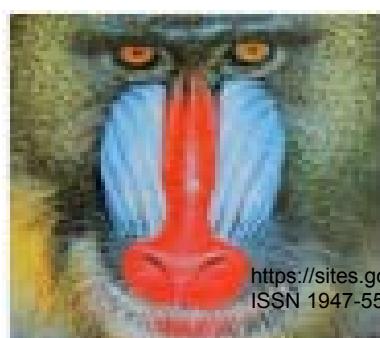
In this way all the plaintexts in the form of bits are combined to make secret data. Now ASCII values are converted into corresponding characters. At the end decrypted data are arranged in matrix form where vector output after decryption of the skew tent map method.

## 4. Experiment and Result Analysis

This section represents the experimental assessment of the proposed Embedding and removal technique for confidentiality of the image. All algorithms and efficacy procedures were implemented using the MATLAB tool on an 4 GB RAM and Windows 7 Professional based 2.27 GHz Intel Core i3 processor.

### 4.1 Dataset

An Experiment completed on the ordinary images such as Baboon, Lena, Boat, Peppers, etc. These are standard images which are derived from <http://sipi.usc.edu/database/?volume=misc>. The system is tested on day to day images as well.

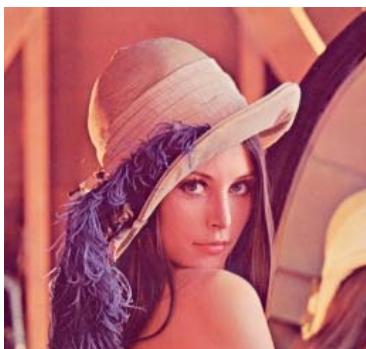


- **Peak Signal to Noise Ratio**

It is a base 10 logarithm ratio of the max pixel value to the mean square error.

$$\text{PSNR} = 10 \log_{10} \left( \frac{\text{Max\_pixel\_value}}{\text{Mean\_Square\_error}} \right)$$

**Fig. 6 Baboon**



**Fig. 7 Lena**



**Fig 8. Boat**



**Fig 9. Peppers**

#### **4.2 Evaluation Parameter**

- **Signal to Noise Ratio**

It is a base 10 logarithm ratio of signal to noise.

$$\text{SNR} = 10 \log_{10} \left( \frac{\text{Signal}}{\text{Noise}} \right)$$

- **Extraction Rate**

It is a percentage ratio of the number of true pixels to total number of pixels present in data hiding.

$$\eta = \frac{n_c}{n_a} \times 100$$

Here  $n_c$  is the number of pixels which are true.

Here  $n_a$  is the total number of pixels present in Data Hiding.

- **Throughput**

It is a ratio of number of completions in bytes to the execution time.

$$\text{Throughput} = \left( \frac{\text{No. of Completions in Bytes}}{\text{Execution Time}} \right)$$

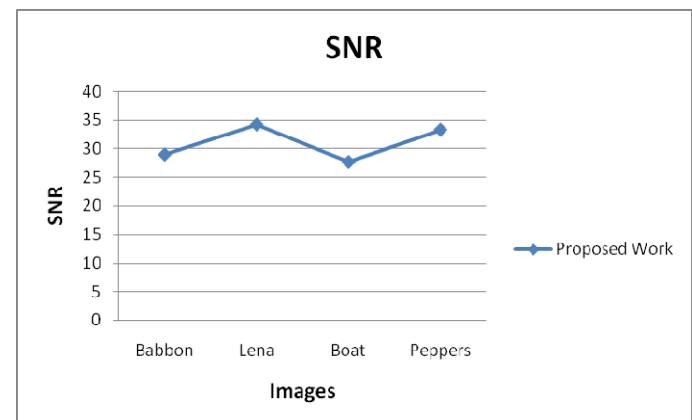
### 4.3 Results Analysis

The execution is obtained in the platform of Matlab with various images. The implementation results are analyzed on the basis of the Signal to Noise Ratio (SNR), Peak Signal to Noise Ratio (PSNR), Data hiding execution time, Data hiding extraction time, Extraction Rate and Throughput. The performance of the proposed algorithm is compared with existing technique on the basis of PSNR, Data hiding execution time and Throughput.

**Table 1. SNR of Proposed Work**

<b>SNR of Proposed Work</b>	
<b>Images</b>	<b>Proposed Work</b>
<b>Baboon</b>	29.0555
<b>Lena</b>	34.2709
<b>Boat</b>	27.7464
<b>Peppers</b>	33.3673

Here, the table1 illustrates that the corollary is attained by means of a Watermarking arrangement in devoid of Noise. In this manner, we acquired the SNR value as 29.0555 through devoid of Noise in Baboon, the SNR value as 34.2709 through devoid of Noise in Lena, the SNR value as 27.7464 through devoid of Noise in Boat, the SNR value as 33.3673 through devoid of Noise in Peppers. Figure 10 represents the SNR values for various images.



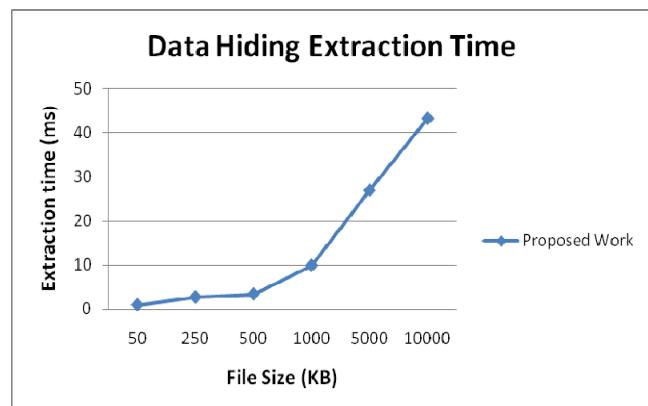
**Fig. 10. SNR of Proposed Work**

**Table 2. Data Hiding Extraction Time of Proposed Work**

File Size (KB)	Data Hiding Extraction Time (ms)
50	0.8974
250	2.6375
500	3.3398
1000	9.8292
5000	26.9362
10000	43.2534

Here, table 2 shows the data hiding extraction time value as 0.8974 for 50 KB size of the file, the data hiding extraction time value as 2.6375 for 250 KB size of the file, the data hiding extraction time value as 3.3398 for 500 KB size of the file, the data hiding extraction time value as 9.8292 for 1000 KB size of the file, the data hiding extraction time value as 26.9362 for 5000 KB size of the file, the data hiding extraction time value as 43.2534 for 10000 KB size of the file. Figure 11 represents the Data hiding extraction time for various file sizes.

as 3.3398 for 500 KB size of the file, the data hiding execution time value as 9.8292 for 1000 KB size of the file, the data hiding execution time value as 26.9362 for 5000 KB size of the file, the data hiding execution time value as 43.2534 for 10000 KB size of the file. Figure 11 represents the Data hiding extraction time for various file sizes.



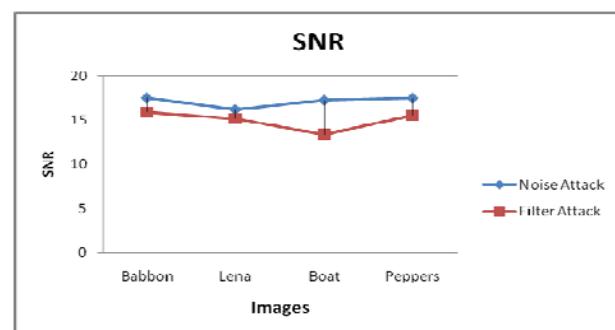
**Fig. 11. Data Hiding Extraction Time of Proposed Work**

**Table 3. SNR of proposed Work under the consideration of Noise and Filter Attack**

SNR		
Images	Noise Attack	Filter Attack
Baboon	17.5420	15.9373
Lena	16.2348	15.2318
Boat	17.2797	13.3664
Peppers	17.5202	15.5973

Here, the table3 illustrates that the corollary is attained by means of a Watermarking arrangement in devoid of Noise. In this manner, we acquired the SNR value as 17.5420 and 15.9373 through devoid of Noise in the baboon, the SNR value as 16.2348 and 15.2318 through devoid of Noise in Lena, the

SNR value as 17.2797 and 13.3664 through devoid of Noise in Boat, the SNR value as 17.5202 and 15.5973 through devoid of Noise in Peppers under the consideration of Noise Attack and Filter Attack respectively. Figure 12 represents the SNR values for various images under the Noise and Filter Attack.



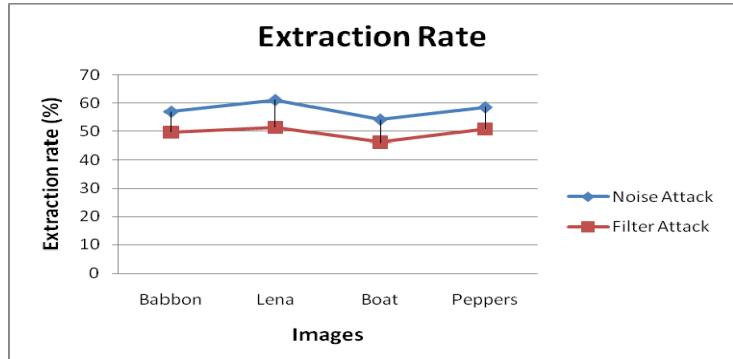
**Fig. 12. SNR of proposed Work under the consideration of Noise and Filter Attack**

**Table 4. Extraction Rate of proposed Work under the consideration of Noise and Filter Attack**

Extraction Rate		
Images	Noise Attack	Filter Attack
Baboon	56.9444	49.7222
Lena	61.1111	51.3889
Boat	54.1667	46.1768
Peppers	58.4948	50.8546

Here, the table4 illustrates that the corollary is attained by means of a Watermarking arrangement in devoid of Noise. In this manner, we acquired the extraction rate as 56.9444 and 49.7222 through devoid of Noise in the baboon, the extraction rate as 61.1111 and 51.3889 through devoid of Noise in Lena, the extraction rate as 54.1667 and 46.1768 through devoid of Noise in Boat, the extraction rate as 58.4948 and 50.8546 through devoid of Noise in Peppers under the consideration of Noise Attack and Filter Attack respectively. Figure 13 represents the extraction rate for various images under the Noise and Filter Attack.

Here, the table5 illustrates that the corollary is attained by means of a Watermarking arrangement in devoid of Noise. In this manner, we acquired the PSNR value as 57.1464 and



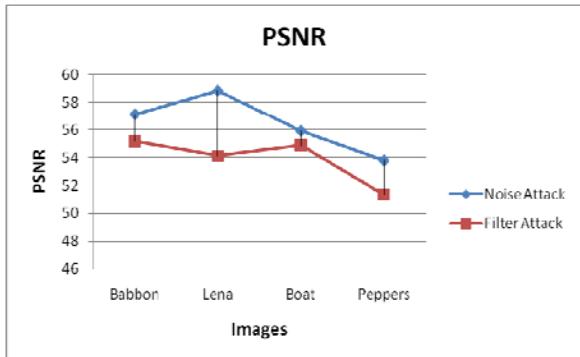
**Fig. 13. Extraction Rate of proposed Work under the consideration of Noise and Filter Attack**

**Table 5. PSNR of proposed Work under the consideration of Noise and Filter Attack**

PSNR		
Images	Noise Attack	Filter Attack
Baboon	57.1464	55.1864
Lena	58.8634	54.1357
Boat	55.9453	54.8743
Peppers	53.8231	51.3633

55.1864 through devoid of Noise in the baboon, the PSNR value as 58.8634 and 54.1357 through devoid of Noise in Lena, the PSNR value as 55.9453 and 54.8743 through devoid of Noise in Boat, the PSNR value as 53.8231 and 51.3633 through devoid of Noise in Peppers under the consideration of Noise Attack and Filter Attack respectively. Figure 14

represents the PSNR values for various images under the Noise and Filter Attack.



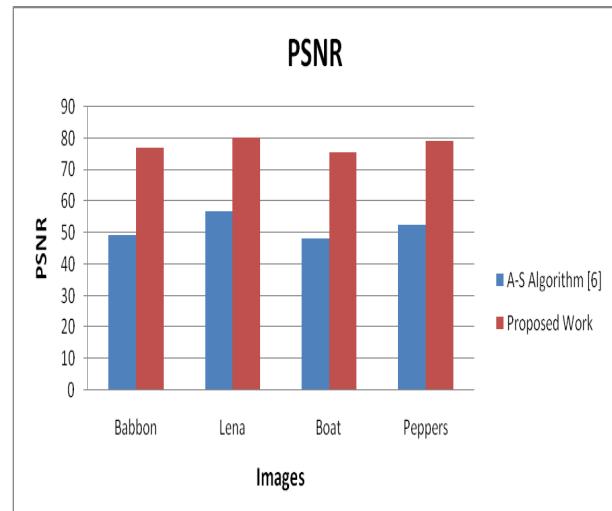
**Fig.14. PSNR of proposed Work under the consideration of Noise and Filter Attack**

PSNR value as 80.4425 through devoid of Noise in Lena, the PSNR value as 75.7464 through devoid of Noise in Boat, the PSNR value as 79.3673 through devoid of Noise in Peppers. From table 6 it is described that under ideal condition proposed work is better than to previous work in [6] under PSNR evaluation parameters. As skew tent and histogram shifting algorithm has regenerate images in color format only so this parameter is high as compare to previous value. Figure 15 represents a comparison of the PSNR values for various images between proposed work and A-S algorithm [6].

**Table 6. PSNR Based Comparison between proposed and previous work (A-S Algorithm) [6].**

Images	A-S Algorithm [6]	Proposed Work
Baboon	49.2	77.0555
Lena	57	80.4425
Boat	48.3	75.7464
Peppers	52.5	79.3673

Here, the table6 illustrates that the corollary is attained by means of a Watermarking arrangement in devoid of Noise. In this manner, we acquired the PSNR value as 77.0555 through devoid of Noise in the baboon, the



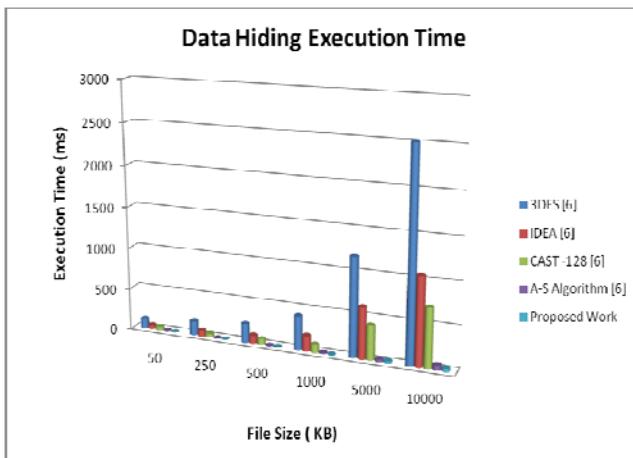
**Fig. 15.PSNR Based Comparison between proposed and previous work (A-S Algorithm) [6]**

**Table 7. Data Hiding Execution Time comparison between proposed and previous works [6].**

File Size (KB)	Data Hiding Execution Time Comparison (ms)				
	3DES [6]	IDEA [6]	CAST -128 [6]	A-S Algorithm [6]	Proposed Work
50	120	49	45	2	0.65409
250	170	69	48	4	1.6352
500	232	101	73	7	4.3348
1000	412	190	96	12	7.8272
5000	1190	621	417	34	25.7161
10000	2507	1059	713	62	41.2718

Here, table 7 shows the data hiding execution time values as 0.65409 for 50 KB size of the file, the data hiding execution time values as 1.6352 for 250 KB size of the file, the data hiding execution time values as 4.3348 for 500 KB size of the file, the data hiding execution

time values as 7.8272 for 1000 KB size of the file, the data hiding execution time values as 25.7161 for 5000 KB size of the file, the data hiding execution time values as 41.2718 for 10000 KB size of the file. From table 7 it is described that under ideal condition proposed work is better as than previous work [6]. As propose work regenerate dictionary from the same data so the execution time for the same is less as compared to previous work. Figure 16 represents a comparison of the Data Hiding Execution Time of various file sizes between proposed work and several previous algorithms [6].



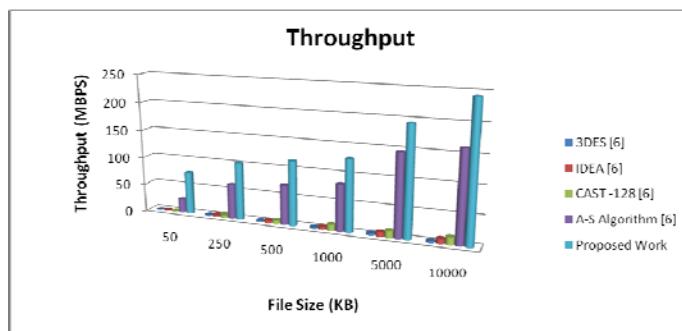
**Fig. 16. Data Hiding Execution Time comparison between the proposed and previous work [6].**

**Table 8. Throughput comparison between proposed and previous works [6].**

File Size (KB)	Throughput Comparison (MBPS)				
	3DES [6]	IDEA [6]	CAST -128 [6]	A-S Algorithm [6]	Proposed Work
50	0.4167	1.0204	1.1111	25	76.4421
250	1.4706	3.6232	5.2083	62.5	102.8865
500	2.1552	4.9505	6.8493	71.4286	115.3456
1000	2.4272	5.2632	10.4167	83.3333	127.7596
5000	4.2017	8.0515	11.9904	147.0588	194.4307
10000	3.9888	9.4429	14.0252	161.2903	242.2962

Here, table 8 shows the Throughput value as 76.4421 for 50 KB size of the file, the Throughput value as 102.8865 for 250 KB size of the file, the Throughput value as 115.3456 for 500 KB size of the file, the Throughput value as 127.7596 for 1000 KB size of the file, the Throughput value as 194.4307 for 5000 KB size of the file, the Throughput value as 242.2962 for 10000 KB size of the file. From table 8 it is described that under ideal

condition proposed work is better than previous work in [6]. As proposed work regenerate dictionary from the same data so the execution time for the same is less than previous work. Figure 17 represents a comparison of the Throughput for various file sizes between proposed work and several previous algorithms [6].



**Fig. 17. Throughput comparison between proposed and previous works [6].**

**References** 1. A. Rasmi, M. Mohanapriya, "An Extensiv Survey of Data Hiding Techniques", European Journal of Applied Sciences, Vol. 9(3), Pp. 133-139, 2017.

2. Abuturab and Muhammad Rafiq, "Multiple color-image fusion and watermarking based on optical interference and wavelet transform", Journal of Optics & Lasers in Engineering, Elsevier, Vol. 89, Pp. 47-58, 2016.

3. Ales Roceka, Karel Slavícek, Otto Dostál, and Michal Javorník, "A new approach to fully-reversible watermarking in medical imaging with breakthrough visibility parameters", Biomedical Signal Processing and Control, Elsevier, Vol. 29, Pp. 44-52, 2016.

## 5. Conclusion

Here, proposed work has efficiently hide data in the carrier image while the security of the carrier is also maintained by encrypting using a skew tent map algorithm. Embedding is done by using a histogram shifting algorithm where pixel value has shuffle as per the secret data. Proposed algorithm will recover or reverse complete data at the receiver end, with carrier image in ideal condition. Results demonstrate that the proposed work is generating the values which retain the image quality and robustness. In the future, work can be enhanced for other attacks and cryptanalysis such as the geometry of the image.

4. Angela Piper, and Reihaneh Safavi-Naini, "Scalable Fragile Data Hiding For Image Authentication", IET Information Security, Vol. 7(4), Pp. 300-311, 2013.

5.. Arun K Mohan, and Saranya M.R, "Improved Reversible Data Hiding Using Histogram Shifting Method", International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES), IEEE Conference, Pp. 1-5, 19-21 Feb. 2015, Kozhikode, India.

6. Aswathy Achuthshankar, Aswin Achuthshankar, Arjun K P, and Sreenarayanan N M, "Encryption of Reversible Data Hiding For Better Visibility and High Security", Global Colloquium in Recent Advancement and Effectual Researches in

- Engineering, Science and Technology (RAEREST), Elsevier, Vol. 25, Pp. 216 – 223, 2016.
7. Akshay Kumar Joshi, Sanjay Sharma, “Reversible data hiding by utilizing AES encryption, image & LZW compression”, International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), Volume 6(2), Pp. 169-176, February 2017.
8. Chen Cui, and Xia-Mu Niu, “A robust DIBR 3D image watermarking algorithm based on histogram shape”, Measurement, Elsevier, Vol. 92, Pp. 130–143, 2016.
9. Chuan Qin, Huili Wang, Xinpeng Zhang, and Xingming Sun, “Self-embedding fragile watermarking based on reference-data interleaving and adaptive selection of embedding mode”, Information Sciences, Elsevier, Vol. 373, Pp. 233–250, 2016.
10. Cui, Chen and Xia-Mu Niu, “A Robust DIBR 3D Image Watermarking Algorithm Based on Histogram Shape”, Elsevier Journal on Measurement, Vol. 92, Pp. 130-143, 2016.
11. Dalel Bouslimia and Gouenou Coatrieux, "A Crypto-Watermarking System for Ensuring Reliability Control and Traceability of Medical Images", Elsevier Journal on Signal Processing, Special issue on Image Communication, Vol. 47, Pp. 160-169, September 2016.
12. Xiaochun Cao, Ling Du, Xingxing Wei, Dan Meng, and Xiaojie Guo, “High Capacity Reversible Data Hiding In Encrypted Images By Patch-Level Sparse Representation”, IEEE Transactions on Cybernetics, Vol 46(15), Pp. 1132-1143, May 2016.
13. Frank Y. Shih and Xin Zhong, "High-capacity multiple regions of interest watermarking for medical images", Elsevier on Information Sciences, [Vol 367–368](#), Pp. 648-659, 2016.
14. Gonge, Sudhanshu Suhas, and Ashok Ghatol, "Aggregation of Discrete Cosine Transform Digital Image Watermarking with Advanced Encryption Standard Technique", Procedia Computer Science Vol. 89, Pp. 732-742, 2016.
15. Hanieh Khalilian, and Ivan V. Bajic, “ Video Data Hiding With Empirical PCA-Based Decoding”, Transactions on Image Processing, [IEEE Signal Processing Society](#), Vol. 22(12), Pp. 4825 - 4840, Dec. 2013.
16. Ioan-Catalin Dragoi, and Dinu Coltuc, “Local-Prediction-Based Difference Expansion Reversible Data Hiding”, Transactions on Image Processing, IEEE, Vol. 23(4),Pp. 1-11, 2014.
17. K. Ma, W. Zhang, N. Yu, and F. Li, “Reversible data hiding in encrypted image by reserving room before encryption”, Trans. Inf. Forensics Security, IEEE, Vol. 8(3), Pp. 553–562, 2013.
18. Kazuki Yamato, Madoka Hasegawa, Yuichi Tanaka and Shigeo Kato, “Digital image

- watermarking method using between-class variance, IEEE, Pp. 1-8, 30<sup>th</sup> Sept.-3<sup>rd</sup> Oct. 2012, Orlando, FL, USA
19. Manas Ranjan Nayak, Joyashree Bag, Souvik Sarkar and Subir Kumar Sarkar, "Hardware implementation of a novel watermarking algorithm based on phase congruency and singular value decomposition technique", International Journal of Electronics and Communication, Elsevier, Vol.71, Pp. 1-8, 2016.
20. Mohan A Chimanna, and S.R.Kho, "Digital Video Data Hiding Techniques For Secure Multimedia Creation And Delivery", IJERA, Vol. 3(2), Pp. 839-844, 2013.
21. Paweł Korus, and Andrzej Dziech, "Efficient Method For Content Reconstruction with Self-Embedding", IEEE Transactions on Image Processing, Vol. 22(3), Pp. 1134 – 1147, 2013.
22. Shahzad Alam, Vipin Kumar, Waseem A Siddiqui and Musheer Ahmad, "Key Dependent Image Steganography Using Edge Detection", Fourth International Conference on Advanced Computing & Communication Technologies, Pp. 85-88, 8-9 Feb. 2014 Rohtak, India..
23. Tamanna Tabassum, and S.M. Mohidul Islam, "A Digital Image Data Hiding Technique based on Identical Frame Extraction In 3-Level DWT", IJARCET, Vol. 13(7), Pp. 560 –576, 2003.
24. Wang, Xiang-yang, Yu-nan Liu, Meng-meng Han and Hong-ying Yang, "Local quaternion PHT based robust color image watermarking algorithm", Elsevier on Journal of Visual Communication and Image Representation, Vol 38, Pp. 678-694, 2016.

#### AUTHORS PROFILE



Anoop Kumar Chaturvedi: obtained his Bachelor's degree from Dr. H.S. Gour University, Sagar, India in

1998 and Master's degree in Computer Science & Engineering with gold medal from Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal, India in 2010. He is currently pursuing his Ph.D. degree in Computer Science & Engineering at Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal, India. His research interest include digital watermarking and image data hiding.



Dr. Piyush Kumar Shukla received his Bachelor's degree in Electronics & Communication Engineering, LNCT, Bhopal in 2001, M.Tech (Computer Science & Engineering) in 2005 from SATI, Vidisha and Ph.D. (Computer Science & Engineering) in 2013 from RGPV, Bhopal. M.P. India. He is a Member of ISTE (Life Member), IEEE, ACM, IACSIT, IAENG. Currently he is working as an Assistant Prof. in Department of Computer Science & Engineering, UIT-RGPV Bhopal. He is also I/C of PG Program (Dual Degree Integrated PG-Programs) in DoCSE, UIT, RGPV, Bhopal, Madhya Pradesh, Bhopal. He has published more than 40 Research Papers in various International & National Journals & Conferences, including 04 Papers in SCIE Journals & More than 10 papers in Scopus Journals.

# Fractal Compression of an AVI Video File using DWT and Particle Swarm Optimization

Shraddha Pandit, Research Scholar

Department of Computer Science & Engineering,  
University Institute of Technology  
Rajiv Gandhi Proudyogiki Vishwavidyalaya  
Bhopal, India  
[svpandit\\_pict@yahoo.co.in](mailto:svpandit_pict@yahoo.co.in)

Piyush Kumar Shukla, Assistant Professor

Department of Computer Science & Engineering  
University Institute of Technology  
Rajiv Gandhi Proudyogiki Vishwavidyalaya  
Bhopal, India  
[pphdwss@gmail.com](mailto:pphdwss@gmail.com)

Akhilesh Tiwari, Associate Professor

Department of Computer Science & Engineering and Information Technology,  
Madhav Institute of Technology & Science  
Gwalior, India  
[atiwari.mits@gmail.com](mailto:atiwari.mits@gmail.com)

**Abstract—** In the current scenario compression of video files is in high demand. Color video compression has become a significant technology to lessen the memory space and to decrease transmission time. Video compression using fractal technique is based on self similarity concept by comparing the range block and domain block. However, its computational complexity is very high. In this paper we presented hybrid video compression technique to compress Audio/Video Interleaved file and overcome the problem of Computational complexity. We implemented Discrete Wavelet Transform and hybrid fractal HV partition technique using Particle Swarm Optimization (called mapping of PSO) for compression of videos. The analysis demonstrate that hybrid technique gives a very good speed up to compress video and achieve Peak Signal to Noise Ratio.

**Keywords-** Video Compression, DWT, Fractal, HV partitioning, MATLAB, MSE, PSO, PSNR, ET, CR.

## I. INTRODUCTION

In current scenario digital video plays very important role in information technology, including [1] teleconferencing, broadcasting, military applications, entertainment and many more [1,2]. People need to access video very quickly and within a limited period of time through various digital devices [3, 4]. To deal with this situation compression of video file [5, 6] is the necessity. The DWT transform [1, 3, 9] form the layers of frames in terms of group of frames [2, 4]. The processing of frames [1, 2, 10] in layer is very slow for the compression [6]. Due to slow compression encoding of video [5, 6, 7] is major problem in DWT based video compression [7, 8]. For the encoding and fast processing [3, 5] transform function [7] used partition process in terms of horizontal and

vertical [6, 8] for the local processing of layers frames in different groups of frames. The reduction of search space [5, 9] in terms of layers of block for coding used PSO (particle swarm optimization) [1, 3]. The PSO [8] reduces the layers space and decrease the encoding time [6] and reduces encoding time bust [5, 10] the performance of video compression [5]. In this proposed work we represent background of discrete wavelet transform [3, 5] and Fractal HV partition technique [5, 7] with PSO [7, 8] in segment II and segment III represent experimental results and lastly segment IV represent conclusion.

## II. BACKGROUND

### DISCRETE WAVELET TRANSFORM

Wavelet transforms [1, 2, 10] is broadly used in computer vision [6, 9] as an image compression [4, 6, 8]. The phenomenon of wavelet is closely allied to multi-scale [2] and multi-resolution [4] application and it has been used into image fusion technique [5, 7]. Implementation of Discrete Wavelet Transform as an image processing method generates the transformation values called wavelet coefficient [5, 7]. The fundamental concept behind wavelets is to examine signal according to scale. During recent years, it has gained a lot of interest in the field of signal processing [5, 7], numerical analysis and mathematics [8]. In general, the wavelet transform is an advanced method of signal and image analysis [4, 5].

## FRACTAL HV PARTITIONING TECHNIQUE AND MAPPING OF PSO

In a HV partition [2, 6, 8] a rectangular range square [8] can be part either on a level plane or vertically into two littler rectangles [4]. A choice about the split area must be made. While embraces a model in view of edge area [4], we take after and propose to part a rectangle with the end goal that an estimation by its DC segment (DC segment of a piece is characterized here as the square whose pixel esteems are equivalent [6] to the normal power of the square.) in each part gives an insignificant aggregate square mistake [5]. We anticipate that fractal coding will deliver moderately little collection mistakes with this decision since approximation by the DC segment [3] alone will as of now give little wholes of squared blunders by plan of the part conspire, and for the guess of the dynamic piece [6] of the range squares we have more areas accessible, if the range piece fluctuations are low [2].

The HV partition technique [3, 7] proceed the video data for the process of encoding in terms of domain and Range block in terms of column for the encoding [4] in terms of horizontal and vertical column of video data[1].

Here in this proposed work to reduce the searching time [6] between range and domain block we used PSO technique with HV partition technique [7] which reduces the searching time of block symmetry and increase the block symmetry.

### III. EXPERIMENTAL RESULTS

In this paper the proposed algorithms of DWT, Fractal transform and PSO algorithm [5] has been implemented using MATLAB 8.0 code. For testing varied audio/video interleaved videos [7], we used a configuration of desktop Intel processor with 1.86GZ with 2 GB of RAM [3] running on Windows 2007. For the evaluation of the performance used some standard parameters such as PSNR, MSE, CR and ET of video [5, 7]. The measured parameter gives better result, instead of DWT based video compression [8]. For testing videos are obtained from CV vision library [6, 8]. All process we describe here.

#### Description of Dataset

Table 1. Shows description of dataset used for compression of varied videos

S.No.	Video Name	Format of video
1	Battle video	Avi
2	Duck video	Avi
3	Cartoonduck video	Avi
4	Lab video	Avi
5	Sumrf video	Avi
6	Gunner video	Avi
7	Airplane video	Avi

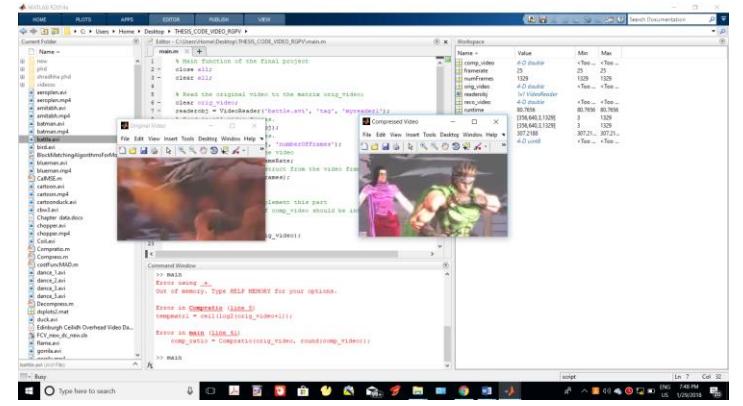


Figure 1. Shows that the original and compressed video view of battle.avi video using DWT method.

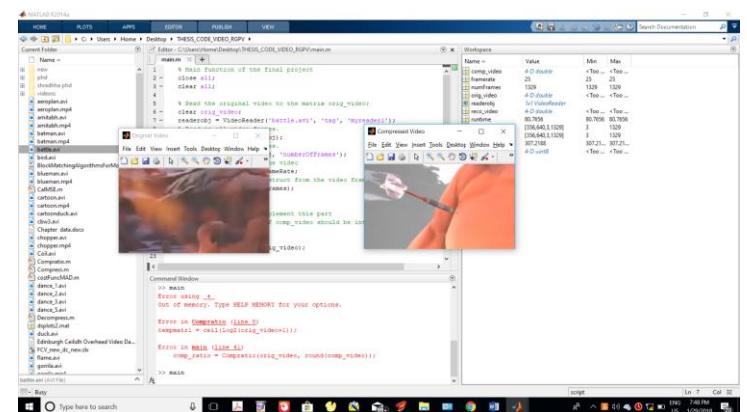


Figure 2. Shows that the original and compressed video view of battle.avi video using mapping of PSO method.

Also get the result of compression of PSNR, Compression Ratio, Mean Square Error and Encoding time for all the tested videos.

The following table shows the comparison of DWT and Mapping of PSO of varied AVI videos with respect to:

- CR-Compression Ratio[8]
- MSE- Mean Square Error Rate[8]
- PSNR- Peak Signal to Noise Ratio[8]
- ET- Estimated Time[8]

Table 2. Analysis for DWT and Mapping of PSO Method for battle.avi video

	DWT	Mapping of PSO
Compression Ratio	0.77	0.81
MSE	11.31	11.09
PSNR	23.14	25.14
Encoding Time	1.80	1.92

Table 3. Analysis for DWT and Mapping of PSO Method for duck.avi video

	DWT	Mapping of PSO
Compression Ratio	0.44	0.59
MSE	11.55	10.21
PSNR	22.34	24.94
Encoding Time	0.52	0.69

Table 4. Analysis for DWT and Mapping of PSO Method for cartoonduck.avi video

	DWT	Mapping of PSO
Compression Ratio	0.89	0.96
MSE	18.74	19.12
PSNR	18.48	21.02
Encoding Time	0.76	0.88

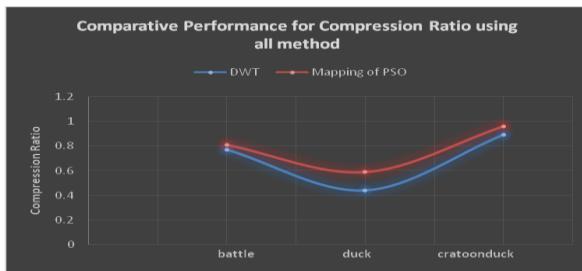


Figure3: Shows the comparative performance of compression ratio using DWT and mapping of PSO method for battle.avi, duck.avi and cartoonduck.avi video

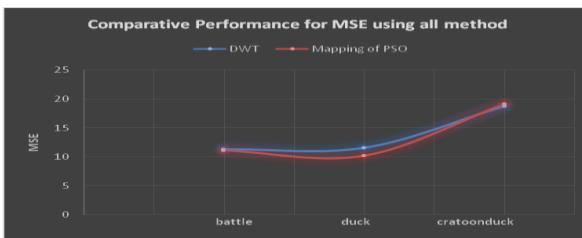


Figure4. Shows the comparative performance of MSE using DWT and mapping of PSO method for battle.avi, duck.avi and cartoonduck.avi video

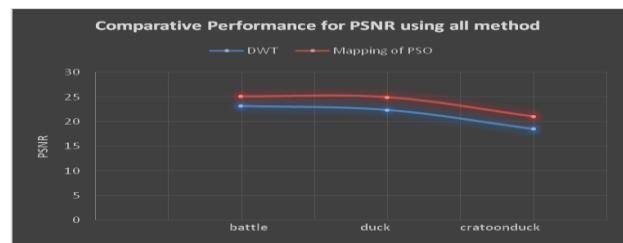


Figure5. Shows the comparative performance of PSNR using DWT and mapping of PSO method for battle.avi, duck.avi and cartoonduck.avi video

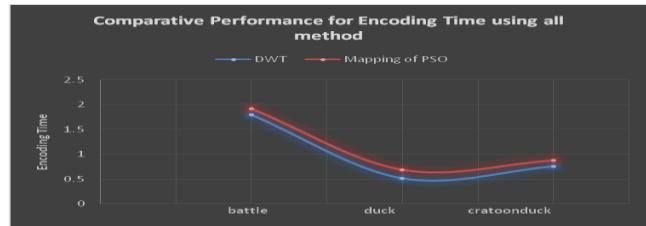


Figure6. Shows the comparative performance of Encoding Time using DWT and mapping of PSO method for battle.avi, duck.avi and cartoonduck.avi video.

#### IV. CONCLUSION

It is essential to reduce the storage space and encoding time of video. From our experimentation and results it is conclude that the DWT transform function faced problem of distortion of layers, due to this reason the value of PSNR is decrease. The particle swarm optimization provides the dual searching mode and reduces the multi-scales H-V partition relation of blocks and references blocks. This reduces space speedup the compression technique and also remains the quality of video. The mapping of PSO also reduces the redundant frames of video and reduces the value of MSE and increase the value of PSNR.

#### ACKNOWLEDGMENT

We express our gratitude towards all the authors which are mentioned in this paper as we are referring their contribution our research.

#### REFERENCES

- [1] Swalpa Kumar Roy,Siddharth Kumar, Bhabatosh Chanda,Biduyt B.Chaudhuri,Soumitra Banerjee, "Fractal image compression using upper bound on scaling parameter", Chaos,Solitons and Fractals, Elsevier, vol.106,Pp.16-22,2018.
- [2] Milind V. Kulkarni, D.B. Kulkarni, "Analysis of fractal inter frame video coding using parallel approach", Journal of Signal, Image and Video Processing, Springer, vol.11, no.4, Pp.629-634, 2017.
- [3] Zhehuang Huang, "Frame-groups based fractal video compression and its parallel implementation in Hadoop cloud computing environment", International Journal of Multidimensional Systems and Signal Processing, Springer, vol.29, no.100, Pp.1-18, 2017.
- [4] Sonali V. Kolekar, Prof.Prachi Sorte, "An Efficient and Secure Fractal Image and Video Compression", International Journal of Innovative Research in Computer and Communication Engineering, vol.4, no.12, Pp.20643-20648, 2016.
- [5] V. Yaswanth Varma,, T. Nalini Prasad, N. V. Phani Sai Kumar, "IMAGE COMPRESSION METHODS BASED ON TRANSFORM CODING AND FRACTAL CODING", International Journal of Engineering Sciences and Research Technology,vol.6,no.10,Pp.481-487,2017.

- [6] Ali Ibrahim Khaleel, N.Anupama, "Haar Wavelet Based Joint Compression Method Using Adaptive Fractal Image Compression", IOSR Journal of Computer Engineering, vol.18, no.3, Pp.69-73, 2016.
- [7] Umesh B Kodgule,B.A.Sonkamble, "Discrete Wavelet Transform based Fractal Image Compression using Parallel Approach", International Journal of Computer Applications,vol.122,no.16,Pp.18-22,2015.
- [8] S.keerthika,S.Vidhya , "Fractal Image Compression with Advanced Particle Swarm Optimization and Sorting",International Journal of Computer Science Trends and Technology,vol.4,no.5, Pp.309-312,2016.
- [9] Gauri R.Desai, Mahesh S.Chavan, "Fractal Image Compression by Range Block Classification", International Research Journal of Engineering and Technology, vol.4, no.1, Pp.525-528, 2017.
- [10] Veena K.K., Bhuvaneswari P, "Various Techniques of Fractal Image Compression – A Review, International Journal of Engineering And Computer Science, vol.4, no.3, Pp.10984-10987, 2015.

#### AUTHORS PROFILE

**Shraddha Pandit** is presently pursuing Ph.D. in Computer Science & Engineering from University Institute of Technology, Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal. She received B.E. in Computer Science & Engineering from Amravati University and M.E. in Computer Science & Engineering from Pune University, respectively. Her research interest include Image processing, Image Compression, Data Mining and Data Warehousing.



**Dr. Piyush Kumar Shukla** received his Bachelor's degree in Electronics & Communication Engineering, LNCT, Bhopal in 2001, M.Tech (Computer Science & Engineering) in 2005 from SATI, Vidisha and Ph.D. (Computer Science & Engineering) in 2013 from RGPV, Bhopal. M.P. India. He is a Member of ISTE (Life Member), IEEE, IACSIT, IAENG. Currently he is working as an Assistant Prof. in Department of Computer Science & Engineering, UIT-RGPV Bhopal. He is also I/C of PG Program (Dual Degree Integrated PG-Programs) in DoCSE, UIT, RGPV, Bhopal, Madhya Pradesh, Bhopal. He has published more than 40 Research Papers in various International & National Journals & Conferences, including 04 Papers in SCIE Journals & More than 10 papers in Scopus Journals.



**Dr. Akhilesh Tiwari** has received Ph.D. degree in Information Technology from Rajiv Gandhi Technological University, Bhopal, India. He is currently working as Associate Professor in the department of CSE & IT, Madhav Institute of Technology & Science (MITS), Gwalior, M.P. (India). His area of current research includes knowledge discovery in databases & data mining, Terrorist Network Mining and Business Intelligence. At present he is acting as a member of Board of Studies at MITS Gwalior, India. He has published more than 50 research papers in the Journals & Conferences of International repute. He is the Editor of Technia: International Journal of Computing Science & Communication Technologies. He is also acting as a reviewer & member in the editorial board of various international journals and program committee member/ member in the advisory board of various international conferences. He is in the panel of experts in various government bodies like UPSC, NIELIT etc. He has guided more than 20 dissertations at Master's level and presently supervising 04 research scholars at doctoral level. He is having the memberships of various Academic/ Scientific societies including IETE, GAMS, IACSIT, and IAENG.



# A Proposal on a Method of ArchiMate based Concept of Operation(ConOps)

Nada Olayan

Graduate School of information science  
Nagoya University  
464-8601 Nagoya, Japan  
olayan.nada@nagoya-u.jp

Shuji Morisaki

Graduate School of information science  
Nagoya University  
464-8601 Nagoya, Japan  
morisaki@is.nagoya-u.ac.jp

Nobuhide Kobayashi

Denso Create Inc.  
460-0008 Nagoya, Japan  
kobayashi.nobuhide@j mbox.nagoya-u.ac.jp

Shuichiro Yamamoto

Graduate School of information science  
Nagoya University  
464-8601 Nagoya, Japan  
yamamotosui@icts.nagoya-u.ac.jp

**Abstract**—There is a strong connection between system success and early identification of users' needs. The Concept of Operation document (ConOps) is an important component for understanding any system clearly on early stages and for guiding a better change process from current state to desired state. There is no specific method that could aid in the process of collecting and analyzing the necessary information to create ConOps document, a lot of analysts face difficulties writing an efficient ConOps document. In this paper we argue that producing ConOps using ArchiMate helps in the process of creating ConOps clauses. We attempt to relate ConOps to ArchiMate Modeling language then prove that ConOps document and ArchiMate modeling language could represent each other efficiently. We also discuss how modeling the ConOps clauses using ArchiMate promotes better understandability and communication to understand the users' needs.

**Key words:** ArchiMate, ConOps, Concept of operation, Enterprise Architecture.

## I. INTRODUCTION

In a previous study, we examined the role of the motivation extension of ArchiMate [1] (ArchiMate is the modeling tool for Enterprise Architecture) and its relation to the adaptiveness of the Enterprise Architecture (EA) [2]. We believe that certain natural language documents could be a factor in the improvement of adaptive EA. Based on the previous examination, we found out that the Concept of Operation document (ConOps) is the ideal document to serve as bases for our research for so many reasons. The IEEE standard1362-1998 (IEEE Std 1362-1998) system definition concept of operation (ConOps)[3] document is widely used for its well-developed guidelines. It is used to guide the transition of systems; it provides a clear description of the current and desired system as well as the description of transition and change. In this research, we start by presenting research's background and research problem in section II. Research methodology and questions in section III.

Next, in order to prove our claim, which states that ArchiMate could aid the process of creating ConOps and vice versa, we first need to establish a connection between ConOps and ArchiMate by mapping elements and then evaluating the mapping and showing a simple demonstration of creating ArchiMate model from ConOps clause by analysis in section IV. In section V we present an experiment to show how using ArchiMate to create ConOps prompts accuracy and reduces difficulty. In section VI, an experimental case study is presented to confirm that using ArchiMate promotes efficient communication and better understanding. Finally, discussion is in section VII and conclusion and future research is in section VIII.

## II. BACKGROUND AND RESEARCH PROBLEM

There are a lot of benefits to creating ConOps document as basis for any system definition, not only do ConOps facilitate communication and consensus among stakeholders, it also play an important role in the whole development lifecycle, because it is used to derive requirements and later used to evaluate the system [11]. After system implementation, the system is validated and verified against the ConOps, which gives a baseline for measuring efficiency [12].

### A. ConOps and the Enterprise Architecture

ConOps IEEE 1362-1998 is the only concept of operation standard that describes current state of system, justifications for change and the concept for the proposed system, which contributes to guiding the transition from current state to future state [3]. This format could serve in tracing and applying changes in a consistent manner. When applying the Architecture Development Methodology (ADM), based on TOGAF 9.1[7], there are three levels of partitioning for enterprise architecture in organizations: strategic level architecture (long term main objective and goal for the organization), segment level architecture (parts of the organizations such as departments) and capability level

architecture. ConOps can be related to each level as shown in Figure 1. Enterprise architecture can be applied to any of the segments (the segment by itself can be identified as an enterprise). As the organization matures each segment would create synergy with the strategic level, this should provide a well-built enterprise architecture that enables flexibility for changes on every level without failure, which leads to creating value and balance to resist failure in the long term.

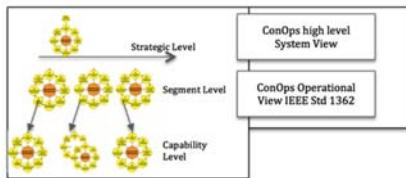


Figure 1. Clarification of Architecture landscape in relation to Conops document

### B. ConOps and ArchiMate Motivation Extension

In this research we utilize ArchiMate to support analysis and problem definition. ArchiMate is an open and independent enterprise architecture modeling language to support the description, analysis and visualization of architecture within and across business domains in an unambiguous way. ArchiMate 3.0 consists of core language, which focuses on the description of the four architecture domains defined by TOGAF standard: business architecture, data architecture, application architecture and technology architecture and two extensions: Motivation and Migration [6]. ConOps main objective is to make sure operational needs are clearly understood by various users. Our hypothesis is that we can improve understandability of the various concepts of the document and help identify operational needs by using ArchiMate concepts and viewpoint. However, detecting motivations can be complex; capturing such concerns in the form of drivers from stakeholders or from the environment can be confusing. That is why we also suggest leveraging ConOps document to get motivation extension aspects and support the architecture transition from present to desired state.

### III. RESEARCH METHODOLOGY

Based on the previous background study we decided to examine the shortcomings of ConOps document. ConOps document is still not used to its fullest potential, in many systems, it's usually made after the system is matured or delivered [4][13]. There are many reasons Conops is underutilized; one of the reasons is the lack of a method for analysis and acquiring the content of the document .The IEEE standard provide guidelines content and format but no exact technique [5]. "Each organization that uses this guide should develop a set of practices and procedures to provide detailed guidance for preparing and updating ConOps documents"[3,iii]. This research is conducted using traditional analysis including background study, comparative and data analysis.

Our hypothesis is that ArchiMate could aid in the process of creating ConOps document. The questions guiding the research are formulated as:

Q1: To which extent can we model ConOps using ArchiMate? (Mapping and evaluation, section 4.1)

Q2: How to produce ArchiMate model from ConOps Clause? (Method and demonstration, section 4.2)

Q3: Does deriving ArchiMate model from ConOps promote understandability and facilitate communication? (Experiment and quantitative analysis, section 5)

Q4: How efficient is it to create ConOps using ArchiMate? (Case study, section 6)

### IV. A PROPOSAL ON USING ARCHIMATE TO CREATE CONOPS

#### A. Mapping ArchiMate Concepts to ConOps clauses

In order to map ConOps clause to the concepts of ArchiMate, we declare classification for ArchiMate concepts such from "Motivation Extension", "Business Layer "and "Application Layer" as shown in Table 1. such as [8]:

- Specialization Problem Element -Meta classes: Stakeholder (S), Driver (D) and assessment (A)
- Specialization Intention -Meta classes (Actual motivation elements): Goal (G), Principle (P), requirement(R), Constraints (Con), outcome (O) and Value (V)
- Strategy Elements: Capability(C), Course of action (CA) and resource(R)

Table 1.Mapping Motivation ArchiMate elements to ConOps clauses

ConOps Element	ConOps Clause	Problem Element			Intention Elements				Strategy Elements				
		S	D	A	G	P	R	Con	O	V	C	CA	R
Scope	1.3 System Overview				✓	✓							
	3.1 Background, Objective and scope	✓	✓	✓	✓	✓							
	3.2 Operational policies and constraints							✓	✓				
Justification and nature of change	4.1 Justification of changes	✓	✓	✓	✓			✓					
	4.2 Description of desired changes	✓							✓	✓	✓	✓	✓
Proposed System	5.1 Background, Objective and scope	✓	✓	✓	✓	✓							
	5.2 Operational policies and constraints						✓	✓					

- "Business Layer" & "Application Layer" elements including all elements in: Behavior entities, Passive entities and Active entities. The mapping is shown in Table 2.

Table 2.Mapping the rest of ArchiMate layers to ConOps clause

ConOps Element	ConOps Clause	ArchiMate Layer
Current System	3.3 Description of the current system or situation	Business Layer
	3.5 user classes and other involved personnel	Application layer
Proposed System	5.3 Description of the Proposed system	Business Layer
	5.5 user classes and other involved personnel	Application layer
	6.0 Operational Scenario	Business Layer

Our efforts to establish connection between ArchiMate and ConOps started by creating ArchiMate Models from a number of ConOps documents. Based on ontology definitions and the created ArchiMate models, we were able to represent coverage of ConOps by ArchiMate as shown in Table 3.

**Table 3.** ArchiMate elements coverage of ConOps clauses

Conops Main parts	Sub-sections	Related ArchiMate element
Scope	1.1 Identification	Out of Scope
	1.2 Document Overview	Out of Scope
	1.3 System Overview	Motivation Extension
Current System	3.1 Background, Objective and scope	Motivation Extension
	3.2 Operational policies and constraints	Motivation Extension
	3.3 Description of the current system or situation	Business Layer, application layer.
	3.4 Modes of operation	Out of Scope
	3.5 user classes and other involved personnel	Business Layer
	3.6 support environment	Out of Scope
Justification and nature of change	4.1 Justification of changes	Motivation Extension
	4.2 Description of desired changes	Strategy Elements
	4.3 priorities among changes	Out of Scope
	4.4 Changes considered but not included	Out of Scope
Concept for proposed system	5.1 Background, Objective and scope	Motivation Extension
	5.2 Operational policies and constraints	Motivation Extension
	5.3 Description of the Proposed system	Business Layer
	5.4 Modes of operation	Out of Scope
	5.5 user classes and other involved personnel	Business Layer
Operational scenario Summary of impact	5.6 support environment	Out of Scope
	6.0 Operational Scenario	Business Layer
	7.0 Summary of Impact	Out of Scope
Analysis of the proposed system	8.0 Analysis of the Proposed System	Out of Scope

### B. Method and demonstration(Producing ArchiMate model from a Conops clause)

To demonstrate modeling Conops using ArchiMate, we used the "United States Government Printing Office (CONOPS V2.0)"[10] as an example. Below is the ArchiMate model for a portion of the Conops clause "1.3 System Overview " (Figure 2).

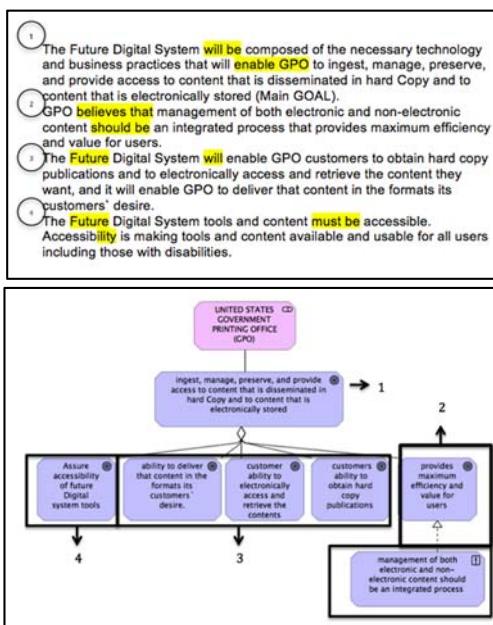


Figure 2. Modeling ArchiMate from ConOps, a) portion form ConOps document (top) and b) the derived ArchiMate model (bottom)

We suggest that the process begins with identifying lead words in each section of the ConOps. "Lead words" are words related to ArchiMate elements based on ontology. Examples of lead words identified include: "future system", "will be",

"Proposed system", "Believe that", "Should be", "primarily", "Services", "However", "as a result", "issue", "failed", "must", etc. After identifying lead words we can match the sentence to the appropriate ArchiMate element based on ontological analysis, in Figure 2.a we show how sentences are divided based on lead words and in 2.b the sentences are assigned to the appropriate ArchiMate element. Deriving the ArchiMate model is based on the analysis shown in Table 4. [8][15]. The recognition of lead words in ConOps could assist in automating the process in the future by creating domain ontologies from textual documents [14].

**Table 4.**Analysis method used to create the corresponding ArchiMate model

ConOps element 1.3System overview		
Lead words	ArchiMate Element	Reasoning (Analysis) (Ontological analysis) from
-Future System.... -Will be... -Proposed system... ...ility.	Goal (Main)	(Proposition) A goal of an undetermined agent. A goal is the propositional content of an agent's intention. From the ME definition, we can observe that: I. A stakeholder is committed to achieving a goal II. Achieving the goal means bringing out certain effects in reality.
-Believe that...Should be...		(Desire which have as content propositions) The propositional content P of the desire is the result of the application of the predicate Q on all systems in a given context, i.e., $P = \forall s (\text{System}(s) \wedge \text{ContextPrinciple}(s)) \rightarrow Q(s))$ where System holds for all systems, Context Principle holds for all systems in the context of application of the principle and Q holds for the systems that exhibit the desired properties stated in the principle.
	Principle	

### V. USING ARCHIMATE TO REPRESENT CONOPS CLAUSES PROMOTES UNDERSTANDABILITY(EXPERIMENT)

In this research, we encourage using ArchiMate to represent ConOps clauses. ArchiMate promotes better understandability and can therefore ease communication and support consensus among stakeholders. Based on the findings from the previous section, we can safely assume that using ArchiMate to model ConOps is valid. But we need to prove that using ArchiMate to model ConOps benefits communication and promotes better understanding than using the ConOps document for stakeholders. For this reason our second assumption is that using ArchiMate models produced from ConOps can promote understandability. To prove our claim we conducted an experiment to measure the level of accuracy and difficulty.

Measuring the understandability would determine if using ArchiMate to represent ConOps is of significant value. For this experiment we had a group of 10 postgraduate students from Yamamoto lab of Nagoya University's department of information science, they were divided in two groups randomly. The participants were familiar with ArchiMate and have had participated in a semester long ArchiMate session in the university. As for their knowledge of ConOps, we introduced the document elements before the experiment from the IEEE Std 1362-1998. For the experiment, we produced a number of ArchiMate models from ConOps (Natural language based) documents then created 2 sets of different questions to test accuracy and difficulty level, The following activities were presented to both groups:

- Activity1: Problem 1(P1) for Group 1(G1): Examined accuracy and understandability using a set of questions based on the original ConOps

(Natural language). Accuracy level: 53%, difficulty level: 2.5

- Activity2: Problem 1 (P1) for group 2(G2): Examined accuracy and understandability using a set of questions based on derived ArchiMate from ConOps (Modeling language). Accuracy level: 93%, difficulty level: 1
- Activity3: Problem 2(P2) for Group 1(G1): Examined accuracy and understandability using a set of questions based on derived ArchiMate from ConOps (Modeling language). Accuracy level: 56%, difficulty level: 2
- Activity3: Problem 2(P2) for Group 1(G2): Examined accuracy and understandability using a set of questions based on the original ConOps (Natural language). Accuracy level: 24%, difficulty level: 3.6

Figure 3.a. shows significant improvement in the level of accuracy when ArchiMate modeling language (ML) is used to represent ConOps as opposed to natural language (NL). Figure 3.b. also shows that difficulty level decreases when participant answered questions from ArchiMate models based on ConOps. Figure 4 shows how Accuracy increases in percentage when levels of difficulty are decreased.

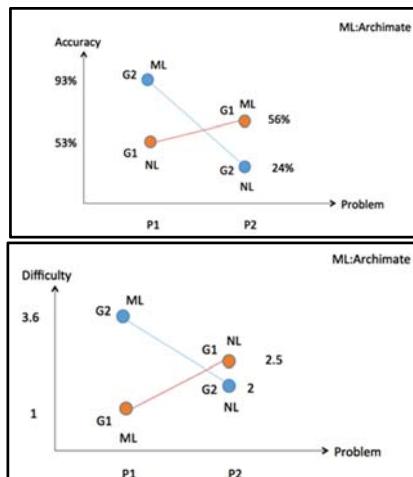


Figure 3. The Effects of natural language and modeling language on "difficulty" and "accuracy" a) shows increase in "accuracy" results (top) b) shows "difficulty" level results (bottom).

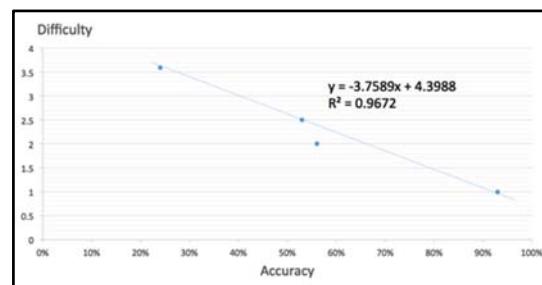


Figure 4. Relationship between accuracy and difficulty in the experiment

## VI. A CASE STUDY ON USING ARCHIMATE CREATE CONOPS CLAUSES FOR "MAN-HOUR MANAGEMENT SYSTEM"

To measure the efficacy of creating ConOps using ArchiMate, we used ArchiMate as the main tool of communication for creating a "Man-hour management" system. Since there is no existing system, we skipped "current system" and included "justification for change" and "description of the proposed system" clauses for this case.

In this experiment we attempt to measure effectiveness of producing ConOps using ArchiMate modeling language. We measured the time of execution for the two main work activities: creating "justification for change" ConOps clause (W1) and creating "description of the proposed system" ConOps clause (W2).

The time for execution included: time for communication, time for creating ArchiMate model and time for creating ConOps from ArchiMate.

The results were compared to the time it took to create the ConOps for the new system without using ArchiMate.

The process of creating ConOps from ArchiMate was far more efficient with better problem definition and analysis, easier communication and less ambiguous concepts and relations' definition.

### A. Case background

The company had problems with productivity of software development. There was a need for improvement, however it was difficult to find the problem. After discussion, it was determined the problem affecting productivity was the lack of a man-hour management system. There was also a concern from the Accounting department, which was related to improving their procedures of payroll management, the procedure involved manual labor that required a lot of time and effort because they relied on paper forms for payment procedures. (Note: ArchiMate Model created by the engineer)

In order to create the corresponding ConOps clauses from the motivation and business model of ArchiMate we needed to have extra communication in addition after creating the models. It took us 20 minutes to create the Motivation model as shown in Figure 5.a. Another 16 minutes were spent on exchanging the extra information as shown in Figure 5.b.

From the ArchiMate model in Figure 5.a, we derived the ConOps clause "Justification for and nature of change" from the motivational model of ArchiMate and from the extra communication as shown in Figure 6.The process of deriving ConOps from the ArchiMate model took 10 minutes.

We then created the Business model in 40 minutes (Figure 7.a), extra communication took 8 minutes (Figure 7.b).

Figure 8 shows the derived ConOps clause "Concepts of the proposed system".

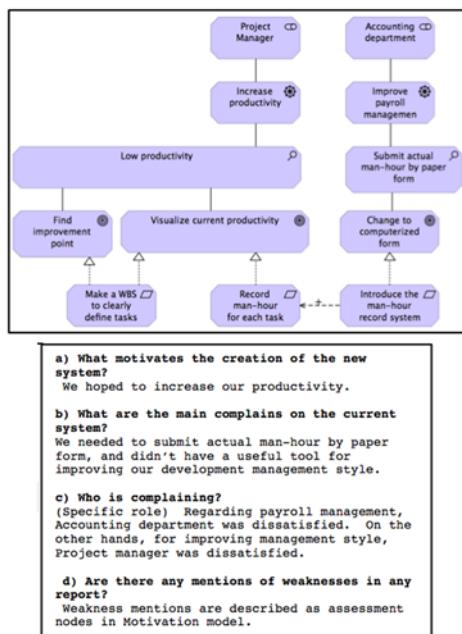
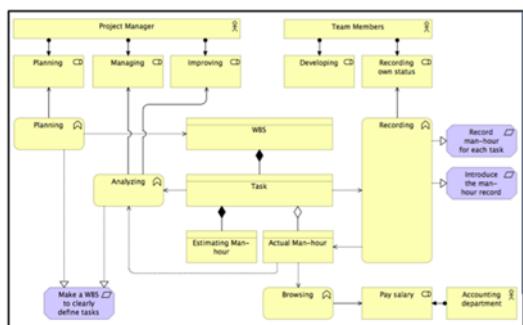


Figure 5. a) Motivation model using ArchiMate (top)  
b) extra Communication (bottom)

**4.0 Justification for and nature of changes (Time 10 minutes)**  
Project manager is concerned about productivity; the current system mainly suffers from low productivity,to address the low productivity issue, visualization of the current productivity and finding improvement points are needed, to achieve the goal of visualization of current productivity there is a need to have a record of man-hour for each task,to find improvement points there is also a need to make work breakdown structure (WBS) to clearly define tasks. There is also a concern raised from the Accounting department. The department wishes to improve payroll management, the analysis shows that the man-hour system is paper based. To improve the payroll management system, based on the previous analysis, change to computerized form is needed.to achieve a fully computerized (automated) system.

Figure 6. the derived ConOps from The ArchiMate model



e) What are the proposed system major components?  
The major component is the project management tool (it was not clearly described in this ArchiMate model). It can provide 4 functions, which were Planning, Analyzing, Recording, and Browsing and the recorded information can be used in payroll management.  
f) What are the main capabilities and functions of the proposed system?  
The main capabilities are increasing productivity, and improving payroll management such as reducing operating time. And functions are Planning, Analyzing, Recording, and Browsing.

Figure 7. a) Business model using ArchiMate (top), b) extra communication (bottom)

**Produced ConOps clause:**  
**5.0 Concepts of the proposed system (Time 15minutes)**  
The proposed new system "project management" will provide the following functions: planning, analyzing, recording and browsing. **"Recording" function:** This function fulfills the requirement "creating the man-hour record" and "Record man-hour". The function serves the role of "recording status" for "team members" stakeholder. Recording function will use information from the object "Task" and send information to "actual man-hour" object. "Task" object as whole is part of "WBS" object and "actual man-hour" is partially part of the object "task". **"Estimating Man-hour" object as whole is part of object "Task".** **"Analyzing" function:** This function fulfills the requirement "make WBS clearly define tasks". It serves the role of "improvement" and "managing" for the stakeholder "project management". The function receives information from the "actual man-hour" and "task" object. **"Planning" function:** This function fulfills the requirement "make WBS clearly define tasks". It serves the role of "planning" for the stakeholder "project management". The function sends results to "WBS" object. **"Browsing" function:** This function serves the role "Pay salary" for the accounting department. It receives data from the object "Actual man-hour".

Figure 8. The derived ConOps from The ArchiMate model

## VII. DISCUSSION

This paper mainly investigates the relation between ArchiMate as a modeling language and ConOps as a natural language based document.

### A. Summary of the findings

In section 3,we formulated the research questions guiding our investigation. Here, we show the main findings (F1, F2, F3, E4) of this research based the research questions:

**F1:** In section 4 we presented a mapping table based on ArchiMate classification for Motivation extension and "Business Layer" and "Application Layer" in table 1 and 2,during our investigation we examined a set of ConOps documents and put the effort to establish connection through modeling all the ConOps clauses using ArchiMate. The results determined showed us the extent of coverage as shown in table 3.

To evaluate the mapping we examined two perspectives of transformation, first one is modeling ArchiMate from ConOps (ConOps perspective), and the second is writing ConOps from ArchiMate (ArchiMate perspective).

For the first part of the evaluation, we examined the coverage from the ConOps perspective and while only 44.4% of ConOps clauses can be represented by ArchiMate elements based on Table.3, we believe that ArchiMate is still suitable tool to facilitate the creation of ConOps document based on the fact that the main challenge of writing the ConOps document stems from the lack of specific analysis methods [5] and some sections considered out of scope for ArchiMate representation in ConOps are mostly descriptive and doesn't require further analysis such as clauses 1.1 Identification, 1.2 Document Overview.

For the second part of evaluation we examine the mapping from ArchiMate perspective (transformation from ArchiMate

to ConOps) we evaluated the mapping based on the evaluation method by [9] figure 9 shows the result of the evaluation. The evaluation shows that mapping is complete, not redundant and there is no excess. However, there is an overload, which means that more than one clause from Conops can be matched to one concept in ArchiMate. The previous two evaluations show that it is more effective to model ArchiMate first to create ConOps document.

Incompleteness	Every ArchiMate concept can be represented by ConOps clauses	Complete
Redundancy	ConOps clauses can be represented by more than one concept in ArchiMate	Not redundant
Excess	All ArchiMate concepts are present in ConOps clause	No Excess
Overload	More than one clause from ConOps can be matched to one concept from ArchiMate	Overload

Figure 9. Evaluation based on mapping ArchiMate elements to ConOps clauses

F2: In section 4.2, we've shown how to model ArchiMate from a ConOps clause by demonstration in figure 2. We had to depend on ontological analysis to extract the corresponding ArchiMate elements because ConOps is a textual document dependent on the author; there are sets of guidelines by the IEEE standard1362-1998, however no specific method for analysis or clear rules. Using corresponding sentences using lead words proved to be useful.

F3: The results of the experiment and quantitative analysis shows increase in accuracy through decrease of difficulty which shows level of understandability of ArchiMate model derived from ConOps was significantly higher, this indicates that a higher level of consonance could also be achieved among stakeholders.

F4: Evaluation Of Case Study: from the previous experimental case study, we conclude that creating ConOps from ArchiMate has many benefits, there were however some shortcomings for creating ConOps from ArchiMate. During communication, some concepts referred to in the ConOps document and their meaning in ArchiMate such as the concept of "Capability" and "Component" mentioned in figure 7 were ambiguous.

If ArchiMate meta-model concepts and relations are specified using OWL-DL specification, then the process of producing ConOps from ArchiMate can be automated and checked for consistency [16]. Process should include using NLP (natural language processing), domain ontologies and ontology models.

Based on the previous findings we assert that ArchiMate can be used to facilitate the analysis process, improve understandability and therefore communication. Other modeling tools have been known to aid in the process of creating ConOps document, using ArchiMate in combination with these tools could promote better coverage.

There have been efforts to describe the advantages of using various systems thinking methods and modeling tools by papers such as [5]. This paper focuses on the relationship between ConOps and ArchiMate specifically and the proposal

of further modeling tools or frameworks to create ConOps is out of this paper's scope.

## VIII. CONCLUSION AND FUTURE WORK

In this paper we investigated the relationship between ConOps and enterprise architecture, we then explained the relationship between ArchiMate concepts and ConOps. Creating ConOps document itself can be challenging because there is no exact technique to collect and analyze the data. For this reason we suggest the use of ArchiMate as a tool for analysis and communication. Based on our findings, we know that using ArchiMate could significantly improve understandability and promote better communication. We presented a mapping table between concepts of ArchiMate and ConOps clauses based on ontology definition, however there is still a need for a complete definition for concepts found in domain specific ConOps documents. To further ease communication and to achieve better connection between ConOps and ArchiMate, we wish to create an accurate definition for most used concepts defined by the ConOps document guidelines and a specific domain ontology and map it to ArchiMate concepts.

Based on our future research activities, we also aim to create a tool based on our definition. The tool should aid in the process of creating a consistent ConOps document and automate the process of identifying concept based on a set of rules using natural language processing and domain ontology and to derive and generate the corresponding ArchiMate models.

## REFERENCES

- [1] The Open Group, ' The Open Group ArchiMate Standard', 1995-2017.[Online].Available:<http://www.opengroup.org/subjectareas/enterprise/archimate>. [Accessed: 20-June- 2017].
- [2] Olayan.Nada,Yamamoto.Shuichiro,"System Thinking Approach to Create CONOPS for the Adaptive Enterprise Architecture", Proceedings of the Asia Pacific Conference on Information Management 2016,pp. 37-39, Hanoi, Vietnam, (20-22 Oct. 2016).
- [3] IEEE Computer Society, March 19, 1998, IEEE Guide for Information Technology—System Definition—Concept of Operations (ConOps) Document (IEEE Std 1362-1998).
- [4] R. Edson, J. Frittman, "Illustrating the Concept of Operations (CONOPs) Continuum and Its Relationship to the Acquisition Lifecycle, "Excerpt from the proceedings of the 7th annual acquisition research symposium., Monterey, CA: Naval Postgraduate School,2010,pp238-257.
- [5] J.Frittman ,R.Edson,"A Systems Thinking Based Approach to Writing Effective Concepts of Operations (ConOps)",8th Conference on Systems Engineering Research, Hoboken, NJ; 17-19 March, 2010.
- [6] Opengroup.org. (2017). What is New in ArchiMate® 3.0 | The Open Group. [Online] Available at: <http://www.opengroup.org/subjectareas/enterprise/archimate/3.0-whats-new-page2> [Accessed 10 Jun. 2017].
- [7] The Open Group, The Open Group Architectural Framework (TOGAF)version 9. feb 2009.
- [8] C. L. Azevedo, J. P. A. Almeida, M. V. Sinderen, D. Quartel, and G. Guizzardi, "An Ontology-Based Semantics for the Motivation Extension to ArchiMate," 2011 IEEE 15th International Enterprise Distributed Object Computing Conference, 2011
- [9] Wand, Y., Weber, R.: On the ontological expressiveness of information systems analysis and design grammars. Information Systems Journal 3(4), 217–237 (1993)

- [10] " UNITED STATES GOVERNMENT PRINTING OFFICE (GPO) CONCEPT OF OPERATIONS (CONOPSV2.0),"for the future digital system/office of innovation and new technology, May16, 2005.Available at: [https://www.gpo.gov/pdfs/fdsys-info/documents/FDsys\\_RD\\_v2.1.pdf](https://www.gpo.gov/pdfs/fdsys-info/documents/FDsys_RD_v2.1.pdf) [Accessed 10 Jun. 2017]
- [11] Office of Management and Budget, December 5, 1994, Operational Concept Description (OCD), Data Item Description DI-IPSC-81430.
- [12] "Concept of Operations," The MITRE Corporation, 10-Apr-2015.Available:<https://www.mitre.org/publications/systems-engineering-guide/se-lifecycle-building-blocks/concept-development/concept-of-operations>.[Accessed: 22-Jun-2017]
- [13] G. Nelson, "The ConOps in a Self-Similar Scale Hierarchy for Systems Engineering", proceedings of Conference on Systems Engineering Research., vol. 69, 2007.
- [14] G. Lame, "Using NLP Techniques to Identify Legal Ontology Components: Concepts and Relations," Law and the Semantic Web Lecture Notes in Computer Science, pp. 169-184, 2005.
- [15] M. Rosemann, P. Green, and M. Indulska, "A Reference Methodology for Conducting Ontological Analyses," Lecture Notes in Computer Science Conceptual Modeling – ER 2004, pp. 110-121, 2004.
- [16] G. C. A. Antunes, M. Bakhshandeh, R. Mayer, J. Borbinha, and A. Caetano, "Ontology-based enterprise architecture model analysis," Proceedings of the 29th Annual ACM Symposium on Applied Computing - SAC 14, 2014

ubiquitous computing, Knowledge creation and Knowledge management. He joined NTT in 1979. He received his B.S. in information engineering from Nagoya Institute of Technology in 1977, and his M.S. and Ph.D. in information engineering from Nagoya University in 1979 and 2000. He joined NTT DATA Corp. in 2002 and had been Deputy Senior Executive Manager (2002-2007). He had also been Research Fellow and Director of Research at the Institute of System Science (2007-2009), Research and Development Headquarters of NTT DATA Corp.

#### AUTHORS PROFILE



**Olayan Nada** received B.S. degree in Computer Science from King Saud University and M.S. degree in Information Science from Nagoya University. She is currently a PhD candidate and a member of Yamamoto lab of Nagoya University where she had done research on dependability assurance. Her current research field is Adaptive Enterprise Architecture.



**Nobuhide Kobayashi** is a senior manager in the research and development department of DENSO CREATE INC.. He received the Ph.D. in information science from Nagoya University, Japan in 2017. He is responsible for introducing methods related to requirements engineering, enterprise architecture and assurance case to actual projects in automotive system development.



**Shuichiro Yamamoto** is a professor of Strategy Office, Information and Communications Headquarters, Nagoya University. Previously, he was engaged in the development of programming languages, CASE tools, network-based smart card environments, and distributed application development platforms. His research interests include distributed information systems, requirement engineering,

# SECURITY LEVEL ANALYSIS OF ACADEMIC INFORMATION SYSTEMS BASED ON STANDARD ISO 27002: 2013 USING SSE-CMM

Endang Kurniawan

Department of Informatics Engineering  
Islamic University of Indonesia  
Yogyakarta, Indonesia  
kurniawan1911@gmail.com

Imam Riadi

Department of Information System  
Ahmad Dahlan University  
Yogyakarta, Indonesia  
imam.riadi@is.ud.ac.id

**Abstract**— this research was conducted to find out the level of information security in organization to give recommendations improvements in information security management at the organization. This research uses the ISO 27002 by involving the entire clause that exists in ISO 27002 check-lists. Based on the analysis results, 13 objective controls and 43 security controls were scattered in 3 clauses of ISO 27002. From the analysis it was concluded that the maturity level of information system security governance was 2.51, which means the level of security is still at level 2 planned and tracked is planned and tracked actively) but is approaching level 3 well defined.

**Keywords:** Academic Information Security; Security System; Maturity Level; ISO 27002; SSE-CMM;

## I. INTRODUCTION

Academic Information System, has been widely used by almost all universities in Indonesia, it is intended to facilitate the delivery of information to learners, and teaching staff and administrative personnel in the management. The more interaction between the system and the user the better system will be vulnerable to being infiltrated or damaged by irresponsible parties. It will be a new issue in terms of security.

Academic information system as a student academic management needs to ensure the security and privacy and integrity of data processed, in addition to the performance of information systems also become an important part that must be considered so that information systems can be used optimally.

The security issues sparked the mechanism to control access to the network in order to protect it from intruders [1]. On software development that supports forensic network is how to determine the appropriate method to facilitate the processing of log data [2]. The system can continue to run in accordance with the needs and their usefulness. It is necessary to process performance measurements taken through examination. In order for an information system security check to work properly a standard is required to do

so. Formally there is no standard reference on what standards will be used or selected by an organization to carry out security checks of information systems, so that it can use standards as required.

Information security is a must. The issue is important because if the information can be accessed by people who are not responsible then the accuracy the information will be doubted can even be misleading information. The following is some formulation of the problems obtained in research whether the system security on the system the academic information used is in accordance with the standards and the extent of the system readiness academic information in the application of information security standards. Besides what role standardize the security of information systems in safeguarding stored information from various threats which exists.

The purpose of this study is to obtain accurate measurement results in terms of information security on academic information systems and improve the quality of information security in accordance with ISO 27002 standard. In addition to knowing the maturity level of security systems used in academic information systems. It is expected that the results can be used as materials considerations in order to prepare measures to improve information security system management

## II. LITERATURE RIVEW

### A. Information Security

Information security is the preservation of information from all possible threats in an attempt to ensure or ensure business continuity, minimize business risk, and maximize or accelerate return on investment and business opportunities [3].

Information security has some aspects that must be understood to be able to implement it. Some of these aspects, the first of three that are most commonly named C.I.A triangle model, as shown in Figure 1[4].



Figure 1. Aspects of Information Security

Confidentiality, integrity and availability are basic requirements for business information security and provide the maintenance requirements of the business [3][4].

- Confidentiality (C): All information must be protected according to the degree of privacy of their content, aimed at limiting its access and used only by the people for whom they are intended;
- Integrity (I): All information must be kept in the same condition in which it was released by its owners, in order to protect it from tampering, whether intentional or accidental
- Availability (A): All the information generated or acquired by an individual or institution should be available to their users at the time they need them for any purpose

It is not limited to data stored in IT systems; all valuable information, regardless of the way it is recorded or stored, needs to be safeguarded. Such data includes not only privacy-sensitive information but also research data and copyrighted materials.

Information security is the protection of information from a wide range of threats in order to ensure business continuity, minimize business risk, and maximize return on investments and business opportunities. Information security is achieved by implementing a suitable set of controls, including policies, processes, procedures, organizational structures and software and hardware functions. These controls need to be established, implemented, monitored, reviewed, and improved, where necessary, to ensure that the specific security and business objectives of the organization are met. This should be done in conjunction with other business management processes. [4]

#### B. SSE-CMM

The Systems Security Engineering Capability Maturity Model (SSE-CMM) was developed with the objective of advancing security engineering as a defined, mature and measurable discipline. The model and its accompanying appraisal method are currently available tools for evaluating the capability of providers of security engineering products, systems, and services as well as for guiding organizations in defining and improving their security engineering practices

SSE-CMM is the Capability Maturity Model (CMM) for System Security Engineering (SSE). CMM is a framework for developing the process, such as the technical process of both formal and informal. SSE-CMM consists of two parts, namely: The Model for process security techniques, projects and organizations, and assessment methods to know the maturity process. The SSE-CMM contains 11 process areas. The definition of each of the process areas below contains a goal for the process area and a set of base processes that support the process area.

- Administer Security Controls
- Assess Impact
- Assess Security Risk
- Assess Threat
- Assess Vulnerability
- Build Assurance Argument
- Coordinate Security
- Monitor System Security Posture
- Provide Security Input
- Specify Security Needs
- Verify and Validate Security

The five Capability Maturity Levels that represent increasing process maturity are:

- Level 0 indicates not all base practices are performed.
- Level 1 indicates all the base practices are performed but informally, meaning that there is no documentation, no standards and is done separately.
- Level 2 planned & tracked which indicates commitment planning process standards.
- Level 3 well defined meaning standard processing has been run in accordance with the definition.
- Level 4 is controlled quantitatively, which means improved quality through monitoring of every process.
- Level 5 is improved constantly indicating the standard has been perfect and the focus to adapt to changes.

SSE-CMM method used by giving the score assessment on each area of the process that selected between 0 to 5 for each process areas[5].

SSE-CMM describes the essential characteristics of the organization's security engineering process which must exist to ensure good security techniques by not advocating the process. Certain or sequential, yet take the general practice observed in the industry.

Assessment and evaluation of investments that have been issued for the implementation of IT is proper to be considered. Based on some research explained that the company has begun to realize and start doing performance measurement and evaluation [6]. In the process of checking the security of information systems, some media used in data storage is very vulnerable to the occurrence of damage that can be done by certain parties [8].

Information security describes efforts to protect computers and non-equipment of computers, facilities, data, and information from abuse by irresponsible people. This definition includes quotes, fax machines, and all types of media, including paper and smartphone documents. For the use of smartphones, in terms of communicating has become a daily necessity. In some cases, smartphone usage may be misused for computer crime, from fraud, to extortion [5].

#### C. Maturity Level

One of the tools of measurement of the performance of system information is a model of maturity level [8]. Maturity model for management and control in the process of information system based on the evaluation methods of the Organization, so that it can evaluate himself from level 0 (none) to level 5 (optimistic). Maturity model is intended to determine the existence of the problem and how to determine the priority of improvement as shown in Table 1.

To identify the extent to which the organization meets the standards information security, can use the identification framework that is represented in a level of maturity that has a level of grouping capabilities of the company.

TABLE I. CRITERIA ASSESSMENT INDEX AT MATURITY LEVEL

Range	Descriptions
0 – 0.50	Non-Existent
0.51 – 1.50	Initial / Ad Hoc
1.51 – 2.50	Repeatable But Infinitive
2.51 – 3.50	Define Process
3.51 – 4.50	Managed and Measurable
4.51 – 5.00	Optimized

Assessment of the ability and maturity of selected IT processes using maturity level, the assessment results show the maturity level of existing IT processes. Next will be determined maturity targets for each selected IT process, the maturity target of each process is the ideal condition that will be achieved in the definition of the desired maturity level, which will then become the reference in the IT management model to be developed.

Once the maturity level of the current process is set and the target of process maturity has been determined, then the gap between the current conditions and the targets to be achieved will be analyzed the identification of opportunities in the gap to be optimized.

Descriptions measurement techniques are made by the nominal size to sort objects from the lowest to the highest, these measurement only give the order rank. Measurements were carried out directly from value that refers to the values of the exiting sorting in maturity models as show in Table 2 [9].

TABLE II. MATURITY LEVEL

Range	Descriptions
0 Existent	The company does not care about the importance of information technology to be managed either by the management
1 Initial	Company reactively performs application and implementation of information technology in accordance with the needs of existing sudden, without preceded by prior planning.
2 Repeatable	The Company has a pattern that is repeatedly performed in conducting activities related to the management of information technology governance, but its existence has not been well defined and that is still happening formal inconsistency.
3 Define	The Company has had formal and written standard operating procedures that have been socialized to all levels of management and employees to be obeyed and worked in daily activities.
4 Manage	The company has had a number of indicators or quantitative measures that serve as targets and objective performance of every application of information technology applications.
5 Optimized	The Company has implemented the information technology governance refers to "best practice"

#### D. ISO 27002

ISO 27002 is published by the International Organization for Standardization (ISO) and the International Electro-technical Commission (IEC). ISO 27002 was originally named ISO/IEC 1779, and published in 2000. It was updated in 2005, when it was accompanied by the newly published ISO 27001. The two standards are intended to be used together, with one complimenting the other.

The standards are updated regularly to incorporate references to other ISO/IEC issued security standards such as ISO/IEC 27000 and ISO/IEC 27005, in addition to add information security best practices that emerged since previous publications. These include the selection, implementation and management of controls based on an organization's unique information security risk environment.

ISO/IEC 27002 is a code of practice - a generic, advisory document, not a formal specification such as ISO/IEC 27001. It recommends information security controls addressing information security control objectives arising from risks to the confidentiality, integrity and availability of information. Organizations that adopt ISO/IEC 27002 must assess their own information risks, clarify their control objectives and apply suitable controls (or indeed other forms of risk treatment) using the standard for guidance.

The standard is structured logically around groups of related security controls. Many controls could have been put in several sections but, to avoid duplication and conflict, they were arbitrarily assigned to one and, in some cases, cross-referenced from elsewhere. For example, a card-access-control system for, say, a computer room or archive/vault is both an access control and a physical control that involves technology plus the associated management or administration and usage procedures and policies. This has resulted in a few oddities (such as section 6.2 on mobile devices and teleworking being part of section 6 on the organization of information security) but it is at least a reasonably comprehensive structure. It may not be perfect but it is good enough on the whole, as shown in Figure 2 [6].

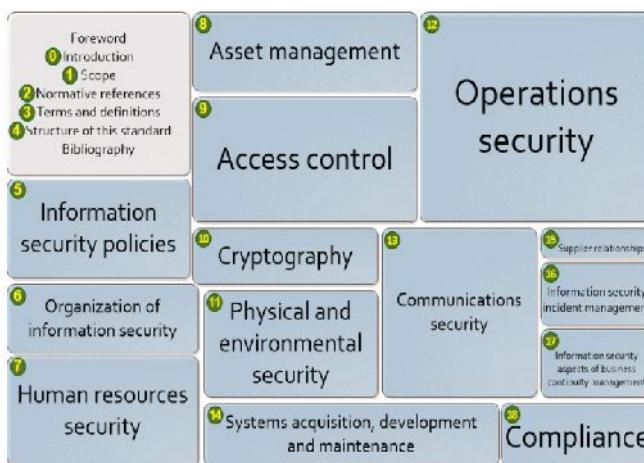


Figure 2. Contents of ISO/IEC 27002:2013

#### E. Gap Analysis

Gap analysis is one tool that can be used to evaluate employee performance. Gap analysis or gap analysts are also one of the most important steps in the planning stages as well as the job evaluation phase.

This method is one of the most common methods used in the management of an institution's internal management. Literally "gap" identifies a disparity between one thing and another. In general, the performance of a company or institution can be reflected in the operational systems and strategies used by the institution. [10]. The service quality gap is defined as the gap between the service that should be provided and the consumer's perception of the actual service provided. The smaller the gap, the better the quality of service

#### F. Guttman Scale

Guttman scaling was developed by Louis Guttman (1944, 1950) and was first used as part of the classic work on the American Soldier.

Guttman scaling is applied to a set of binary questions answered by a set of subjects. The goal of the analysis is to derive a single dimension that can be used to position both

the questions and the subjects. The position of the questions and subjects on the dimension can then be used to give them a numerical value. Guttman scaling is used in social psychology and in education.[7]

The Guttman scale is one of the three major types of unidimensional measurement scales. The other two are the Likert Scale and the Thurstone Scale. A unidimensional measurement scale has only one ("uni") dimension. In other words, it can be represented by a number range, like 0 to 100 lbs or "Depressed from a scale of 1 to 10". By giving the test, a numerical value can be placed on a topic or factor.

The scale has YES/NO answers to a set of questions that increase in specificity. The idea is that a person will get to a certain point and then stop. For example, on a 5-point quiz, if a person gets to question 3 and then stops, it implies they do not agree with questions 4 and 5. If one person stops at 3, another at 1, and another at 5, the three people can be ranked along a continuum.

### III. RESEARCH METHODS

This chapter describes how research where there are details about the material or the materials, tools, sequence of steps to be made in a systematic, logical so it can be used as guidelines are clear and easy to resolve the problems, analysis of results and the difficulties encountered. The sequence of steps problem solving research can be seen in Figure 3.

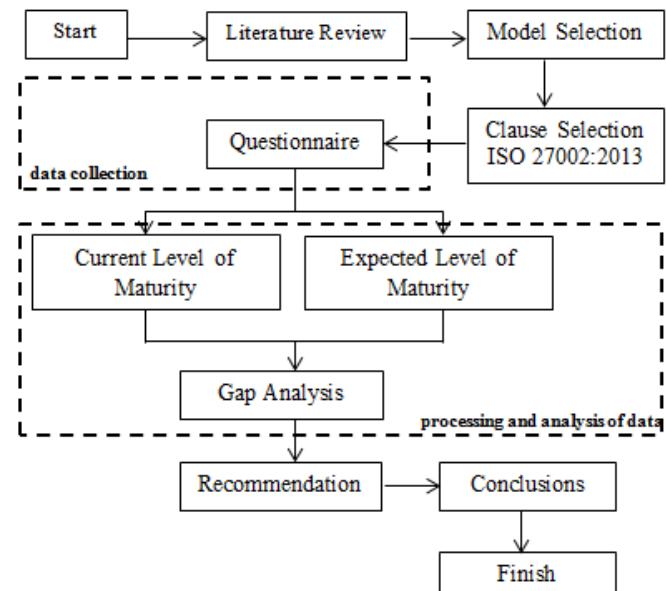


Figure 3. Steps of Research Activities

In this research, the method used is qualitative research method, which data obtained based on the results of questionnaires distributed to respondents. In distributing questionnaires the authors make a list of questions based on the standards contained in ISO 27002 on instructions implementation of information security management which consists of 3 criteria or clauses. The scope of security checks

of information systems is done by determining the control objectives to be used. And it has been engagement letter of agreement that has been made before. Company needs to make an election to the existing controls by taking into account its organizational needs, how to apply and determine risks if those controls are not met. Controls are designed to provide assurance that managerial actions can ensure that business objectives will be achieved and unwanted events will be prevented, detected and improved (Sarno, 2009) [12]. Table 3 is a mapping of the guidelines used against ISO 27002 clauses.

TABLE III. CLAUSE OF ISO 27002

Clauses	Descriptions
9	Access Control
11	Physical and Environmental Security
14	System Acquisition, Development, and Maintenance

Secondary data that the authors use in this study obtained through literature or literature studies such as books, journals, and proceedings. From the results of questionnaires spread then processed by using Maturity Level to get the results of the calculation level of information security maturity. The scale used in this questionnaire uses the Guttman scale. Measurement scale with this type, will get a firm answer, that is yes-no, right-wrong, never-never, positive-negative and others.

In this research answer questionnaire provided two choices that are choice Yes and answer No. In the calculation, the answer Y (Yes) is converted to a value of 1, and the answer N (No) is converted to a value of 0.

The software used in this maturity level calculation is Microsoft Excel. After all the results of the questionnaire are included in the table, then calculated maturity level of each process in each clause for each respondent. Based on the questionnaire that was distributed to respondents selected for filling out the questionnaire in this study were 7 respondents, as shown in Table 4.

Analysis and interpretation of data from the results of data processing and interviews with the manager of academic information systems can be used as research findings, based on the results of maturity level calculation, it can see the gap and can determine the expected value that will make recommendations from each control objective that need improvement.

TABLE IV. RESPONDENTS

No	Functional Structure	
1	Head of Information Technology	1
2	Assistant of information technology Development of Systems and Applications	1
3	Executing Administration Information System	2
4	Senior Executing Data Processing and Reports	1
5	Programmer	2
<b>Respondents</b>		<b>7</b>

#### IV. RESULT AND ANALYSIS

In this section will explain the results of analysis on the implementation and performance measurement of the maturity level of academic information system security obtained from the results of questionnaires and interviews in accordance with the framework of ISO / IEC 27002.

##### A. Summary of the Maturity Level

Based on the results of the recapitulation of the results of questionnaires spread then made an average answer to the questionnaire calculated based on clauses and respondents to get the maturity level, the results are as follows:

###### 1) Maturity level Result Clause 9 : Access Control

Based on the calculation of maturity level, the value obtained in clause 9 about access control information security is at the Initial / Ad Hoc level on position value of 1.44 which means current information security academic information system yet in accordance with standard processes and should be improved. Duties and responsibilities information security should be implemented by all staff who run the information system academic. Third parties are not allowed to access non-information is authorized, third parties may only access general data, as shown in Table 5.

TABLE V. CALCULATION OF CLAUSE 9: ACCESS CONTROL

Control Object.	Description	Index
9.1.1	Access control policy	1.54
9.2.1	User registration	1.10
9.2.2	Privilege or special management	0.89
9.2.3	User password management	0.50
9.2.4	Review of user permissions	1.20
9.3.1	Use of passwords	0.90
9.3.2	Unattended user tools	2.10
9.3.3	Clear desk and clear screen policies	1.20
9.4.1	Network service usage policy	1.40
9.4.2	User authentication to connect out	1.00
9.4.5	Separation with the network	1.20
9.4.6	Control over network connections	0.50
9.4.7	Control of network routing	1.20
9.5.1	Safe log-on procedures	1.67
9.5.2	User identification and authentication	2.67
9.5.3	Password management system	2.78
9.5.4	Use of system utilities	1.50
9.5.5	Session time-out	1.40
9.5.6	Connection timeout	1.75
9.6.1	Information access restrictions	1.90
9.6.2	Isolate sensitive systems	1.20
9.7.1	Communication and computerized moving	2.00

The maturity level calculation results in clause 9 can be represented in graphical form. The result of the maturity level calculation of clause 9 of access control can be seen in Figure 4.

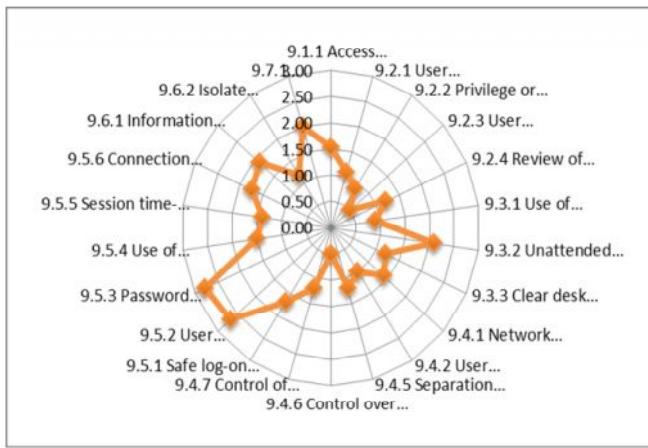


Figure 4. Maturity Level of Clause 9: Access Control

### 2) Maturity level Result Clause 11 : Physical and Environmental Security

Based on the calculation of maturity level value obtained in process 11 about physical and environmental security is at the level of repeatable but intuitive at position value of 2.47 which means current information security academic information system should be developed into better stages. Until now there has been no security audit process information on academic information systems, but the policy issued by management is evenly distributed to all existing parts. Important notes or important information is protected by the system to avoid damage and loss, as shown in Table 6.

TABLE VI. CALCULATION OF CLAUSE 11: PHYSICAL AND ENVIRONMENTAL SECURITY

Control Object.	Description	Index
11.1.1	Physical security restrictions	4.14
11.1.2	Physical in control	2.1
11.1.3	Security office, space, and amenities	3.23
11.1.4	Protection against external attacks and environmental threats	1.1
11.1.5	Working in a safe area	3
11.1.6	Public access, shipping area and drop of goods	3.5
11.2.1	Placement of equipment and protection	4.3
11.2.2	Supporting utilities	2.34
11.2.3	Security of wiring	1.71
11.2.4	Equipment maintenance	1.56
11.2.5	Safety equipment outside the workplace that is not hinted	2.41
11.2.6	Security of disposal or re-use of equipment	2.5
11.2.7	Right transfer of equipment	0.25

The maturity level calculation results in clause 11: physical and environmental security can be represented in graphical form, as show in Figure 5

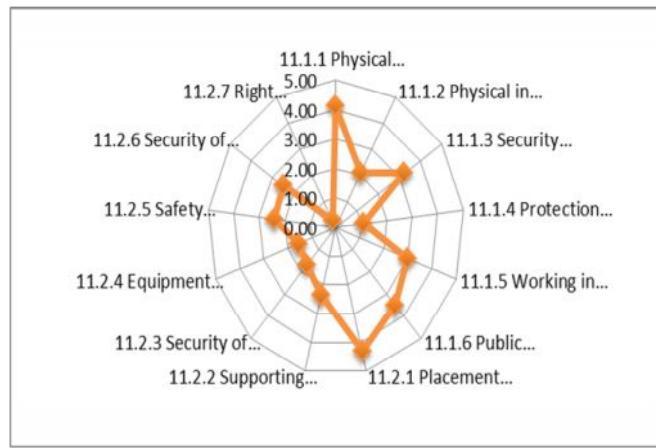


Figure 5. Maturity Level of Clause 11: Physical and Environmental Security

### 3) Maturity level Result Clause 14 : System Acquisition, Development, and Maintenance

Based on the calculation of maturity level value obtained in process 14 about the acquisition of information, development and maintenance systems is at the managed and measureable level at position value of 3.63 which means information security is standard and must documented and then published through training. Information systems academic is an interactive system because every validation, the system will be issues messages related to user-initiated activities. All information systems designed and built by the Information Technology Division without any interference hands of outsiders and out sourcing, as shown in Table 7

TABLE VII. CALCULATION OF CLAUSE 14: SYSTEM ACQUISITION, DEVELOPMENT, AND MAINTENANCE

Control Object.	Description	Index
14.1.1	Incorporate information security in the business continuity management process	3.85
14.2.1	Validate input data	3.5
14.2.2	Controls for internal processing	3.9
14.2.4	Validation of output data	3.4
14.5.1	Additional control procedures	3.68
14.5.3	Restrictions on software package changes	3.51
14.5.4	Weakness of information	3.42
14.6.1	Control of technical weakness (Vulnerability)	3.75

The maturity level calculation results in clause 14: system acquisition, development, and maintenance can be represented in graphical form. The result of the maturity level calculation of clause 11 can be seen in Figure 6.

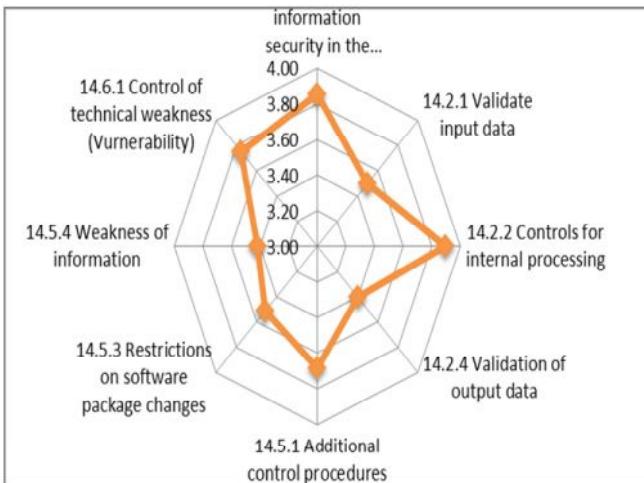


Figure 6. Maturity Level of Clause 14: System Acquisition, Development, And Maintenance

The value of maturity obtained from the average of respondents' answers to each clause contained in the ISO 27002 standard. Table 8 shows the results of the calculation of the questionnaire to obtain the level of maturity academic information system.

TABLE VIII. RESULT OF CALCULATION MATURITY

Clauses	Descriptions	Index	Level
9	Access Control	1.44	1
11	Physical and Environmental Security	2.47	2
14	System Acquisition, Development, and Maintenance	3.63	3
<b>Average maturity level</b>		<b>2.51</b>	<b>3</b>

The result of the calculation to get the average value of information security control on the academic information system of 2.51 From this value, it can be concluded that the security information is at level three, which is well defined or mean standard processing has been run in accordance with the definition.

Based on the result from Table 8, for each process in the clause, it obtained graphs as in the Figure 7 below.

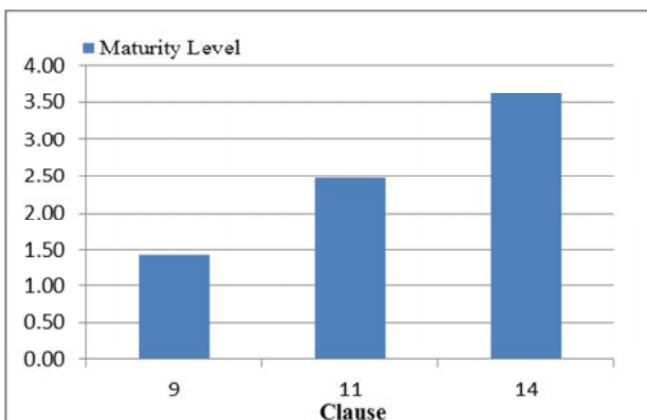


Figure 7. Measurements graphs in maturity level

### B. The Gap of Value Maturity Level

Based on the calculation of information security maturity level of the current academic information system in 2.51 (define) and expected the maturity 1 level is 5 (optimized). The reason is its value achieved by level five (optimized) is the organization's readiness in the field security policies, procedures and processes, and access control information security, can be seen Table 9 below.

TABLE IX. RESULT OF CALCULATION MATURITY

Clause	Description	Maturity		Gap
		Current	Expected	
9	Access Control	1.44	5	3.56
11	Physical and Environmental Security	2.47	5	2.53
14	System Acquisition, Development, and Maintenance	3.63	5	1.37
<b>Average</b>				<b>2.49</b>

Based on Table 9, the distance gap between the current conditions with the expected conditions for each clause is a clause 9 values gap value of 3.56, clause 11 value gaps of 2.53, and in Clause 14 gap value is 1.37.

After getting the value gap for each clause then all values are summed gap then averaged to obtain the value of the overall gap. The overall value of the gap there is a distance of 2.49 between the maturities of the current conditions with the maturity of the expected conditions. There is the fairly large gap, then the required adjustment of each control. Recommendations will be given to each control so much focus on the improvement of weak controls. Value ratio of the current maturity level and the value of the expected maturity level are depicted in Figure 8.

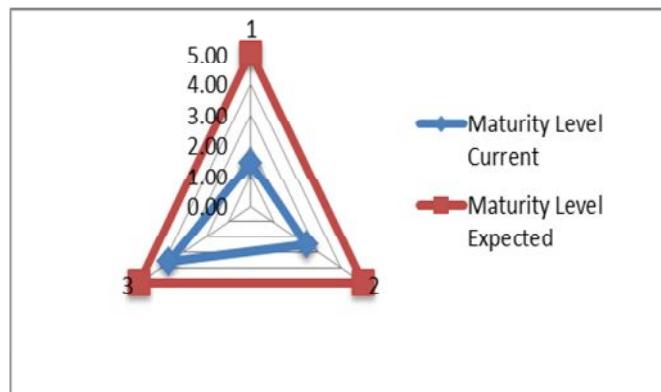


Figure 8. Result of Gap Analysis

As shown in figure 8 that the current condition of the current Maturity level (current level) is represented on the blue line while the expected Maturity level (target level) is on the red line. It is seen that the expected maturity level of five is improved continuously which signifies the standard has been perfect and the focus to adapt to change. Level selection these targets are based on the consideration of the

results of the analysis where the objective value of the control is spread across the range values 1 and 3.

And it is explained that the current level of security of the lowest value of the analysis gap is 3.56 in clause 9 with the maturity level of information security at the level of 1.44 current conditions. While the highest maturity value in clause 14 with the value of maturity level reached 3.63 so it has the lowest analysis gap value is 1.37. Thus the higher the value gap clause, the more likely the clause is to get a security breach and the lower value of the gap in clause then the less likely the clause is to get security problems.

## V. RECOMMENDATIONS AND IMPROVEMENT STRATEGY

After doing analysis and evaluation of information security, the researcher got some conditions in accordance with the ISO 27002 security controls have been set. Some of these conditions are:

- There are rules about information security responsibilities in the employment contract employees.
- There is a perimeter security to protect areas which contain information processing facilities.
- The determination of business requirements for access control.
- There is a responsibility of management on information security incident management.

While conditions still need to be improved are:

- The confidentiality agreement has not been described in detail and specifics.
- There are no training related to information security, such as the criteria for good passwords, training in anticipation of a virus attack..
- Do not do the review and reform of the right of access on a regular basis. Renewal of permission is not required on a regular basis.
- There are many policies and procedures have not been documented even some action in organization conducted by spontaneity and without any irregular formal.
- Reexamination of the access rights of each and updating access rights in case of transfer of part or advancement in accordance with their respective access rights.
- Every employee, contractor or third party should return all the company's assets used for work depending on the contract, when the employee, contractor or third party quit the company or moved other part.

Defects equipment information system is one of system that occurs due to lack of maintenance carried out by the organization, a lack of management capacity and handling equipment made less coordinated.

## VI. CONCLUSIONS

Based on the result of security analysis of information system in this research, determined 13 objective control and 43 security control spread in 3 clause of ISO 27002 used in process of information security system audit.

The SSE CMM measures maturity levels of the relevant security processes that an organization implements to achieve tended capability maturity levels of the security Pas. Application of the SSE-CMM is a straight forward analysis of existing processes to determine which base processes have been met and the maturity levels they have achieved. The same process can help an organization determine which security engineering processes they may need but do not currently have in practice

The results obtained from the measurement of the level of maturity for academic system information is level 3 (well define). Results of the questionnaire management to obtain an average value for all of the clauses is 2:51 range of 0 to 5. And the value of the gap between current security conditions and the condition of the expected 2.49. From this value can be concluded that the security information on the level three, is defined process.

Thus the results level of security is at level define process means the standard process has been running in accordance with the definition or in other words based on the vision, mission, the objectives, and direction of the organization's development procedures are standardized and documented and communicated through training, but the implementation is left to the team to follow the process, so that the deviations are known, the procedures are refined for the formalities of existing practice

## REFERENCES

- [1] Hermaduanti, Ninki & Riadi, Imam. (2016). Automation framework for rogue access point mitigation in ieee 802.1X-based WLAN. Journal of Theoretical and Applied Information Technology. 93. 287-296
- [2] Imam Riadi, Jazi Eko Istiyanto, Ahmad Ashari and Subanar, "Internet Forensics Framework Based-on Clustering" International Journal of Advanced Computer Science and Applications(IJACSA), 4(12), 2013. <http://dx.doi.org/10.14569/IJACSA.2013.041217>
- [3] S. Faris, H. Medromi, S. El Hasnaoui, H. Igner and A. Sayouti, | Towards an Effective Information Security Risk Management of Universities Information Systems Using Multi Agent System, Itil, ISO 27002, ISO 27005 (IJACSA) Internasional Journal of Advanced Computer Science and Application, Vol. 5 No. 6 2014, pp 114 – 118.
- [4] Ermana, F. H., Tanuwijaya Mastan, I. | Security audit information system based on the ISO 27001 Standards on PT. BPR Jatim, STIKOM, Surabaya, 2012
- [5] Adi Supriyatna. | Analysis of the academic information system security level by combining Standard BS-7799 with SSE-CMM, Prosiding Seminar Nasional Aplikasi Sains & Teknologi (SNAST), ISSN: 1979-911X, Yogyakarta, November 2014
- [6] Riyanto Sarno and Irsyati Iffano, | Information Security Manajemen System, Surabaya: ITSPress 2009 (in Indonesian Language).
- [7] Abdi, H., & Williams, L.J (2010). Correspondence analysis. In N.J. Salkind (Ed.): Encyclopedia of research design. Thousand Oaks (CA): Sage.

- [8] Rosmiati & Riadi, Imam & Prayudi, Yudi. (2016). A Maturity Level Framework for Measurement of Information Security Performance. International Journal of Computer Applications. 141. 975-8887. 10.5120/ijca2016907930
- [9] Widiyasono, Nur & Riadi, Imam & Luthfi, Ahmad. (2016). Investigation on the Services of Private Cloud Computing by Using ADAM Method. International Journal of Electrical and Computer Engineering. 6. 2387-2395. 10.11591/jece.v6i5.11527.
- [10] Kohar, Abdul & Riadi, Imam & Lutfi, Ahmad. (2015). Analysis of Smartphone Users Awareness Activities Cybercrime. International Journal of Computer Applications. 129. 1-6. 10.5120/ijca2015906449
- [11] Jackson, Marry Jo, et al. 2011. Quality as a gap analysis of college student's expectation. Jurnal Internasional. USA
- [12] Herison Surbakti, Cobit 4.1 A maturity Level Framework for Measurement of Information System Performance (Case Study : Academic Bureau at Universitas Respati Yogyakarta, International Journal of Engineering Research & Technology (IJERT), Vol. 3, Agustus 2014, ISSN:2278-0181, pp 999 – 1004.
- [13] Surni Eniwati and Nina Kurnia Hikmawati, | An Analysis of Information Technology on Data Processing by using Cobit Framework, (IJACSA) Intermasional Journal of Advanced Computer Science and Application, Vol. 6 No. 9 2015, pp 151 - 157.
- [14] Sarno, R. and Iffano, I. 2009. Information Security Management System. Surabaya: ITS Press.
- [15] Sugiyono. (2007). Quantitative Research Methods, Qualitative and R & D. Bandung. CV.Alfabeta.
- [16] Sandjaja, B & Albertus Heriyanto. (2006). Research Guide. Jakarta. Prestasi Pustaka Publisher

# Simulation and coding of a neural network, performing generalized function with real inputs

Elitsa Zdravkova Spasova

Faculty of Mathematics and Informatics  
University of Shumen "Episkop Konstantin Preslavsky"  
9700, Shumen, Bulgaria  
e.zdravkova@shu.bg

**Abstract** - This paper describes a realization and research on a neural network for a generalized function with real inputs and a binary output (0 or 1). The neural network has been implemented of three different tools - a neural network simulator NeuroPh, logic programming language Visual Prolog and object - oriented programming language Java. The aim is to explore the neural network realization capabilities of the three tools - a neural network simulator, a logical programming environment, and a language for object-oriented programming. For this purpose is selected function with real inputs and binary output (0 or 1) whose values the neural network is trained to predict. The results obtained allow identifying the strengths and weaknesses of the three realized neural networks as well as the environments through which they are realized and tested.

**Keywords** - neural networks, simulators, logic programming language, Visual Prolog, Java

## I. INTRODUCTION

There are various methods and algorithms for modeling artificial neural networks (Nachev & al., 2009, 2011). [5,6]

The simulators are one of these methods, eg NeuroPh and Joone. They have a number of advantages as a convenient interface and tools to easily build models of various types of neural networks (Zdravkova & Nenkov, 2016 a, b). They are available free environments with GNU Lesser General Public License (LGPL) and easy to absorb.

The simulator selected for implementation to the neural network in this study is NeuroPh, which is Java - based, object - oriented simulator. NeuroPh is also open-source and it offers many opportunities for different architectures of neural networks. NeuroPh is lightweight frameworks allowed to simulate neural networks and can be used basic for the development of standard types of neural network architectures. It contains well designed open source library and a small number of core classes that correspond to basic concepts in neural networks. There is a good graphics editor to quickly build java - based components of neural networks.

Neural networks for its remarkable ability to derive meaning from complicated and inaccurate data can be used to extract patterns and detect trends that are too complex to be noticed by humans or other computer techniques. Neural network learning can be considered as "experts" in the field of information, which is given for analysis. These experts can be used to make predictions qualitatively new situations and answer questions such as "what if".

Neural networks take a different approach to problem solving than that of conventional computers. Conventional computers use an algorithmic approach ie computer follows a set of instructions for solving problems. Computer can not solve the problem if it is not aware of specific steps that follow. This limits the ability to solve problems conventional computers to problems that are already known and have a solution. But computers would be much more - useful if they could do things that people do not know exactly how to do.

Neural networks process information in a way similar to the human brain. The network consists of a large number of highly interconnected elements that work in parallel to solve a specific problem. Neural networks learn by example. They can be programmed to perform a specific task. Examples must be selected carefully, otherwise lost valuable time or even worse the network may not function properly. The downside is that since the network "reveals itself" how to solve the problem, its actions can not be predicted. [1,2,4]

The function selected for the survey is with input data, which are real numbers. Such data is also obtained by using inputs from sensors and sensors.

## II. METHODOLOGY

In order to effectively test neural network implementations, it is necessary to create a more complex, generalized function. It will be implemented through a neural network of the NeuroPh simulator, the Visual Prolog logic development environment, and the object-oriented Java programming language. Through the test results, the link between the selected neural network encoding method and its performance can be seen. It will be possible to see again the advantages and disadvantages

of each of the environments in the realization of a neural network, and how this influences its learning and results.

The selected function has two inputs, which are real numbers and one output is an integer (0 or 1). The number used for training is shown in Table. 1.

Implementation of the NeuroPh Network Simulator is required. It is a Java-based simulator that has a good interface and a variety of training and neural network building options. In this case, the neural network needs two inputs and one output. The function is not typically linear, so a hidden layer is also needed. The architecture of the neural network looks like following:

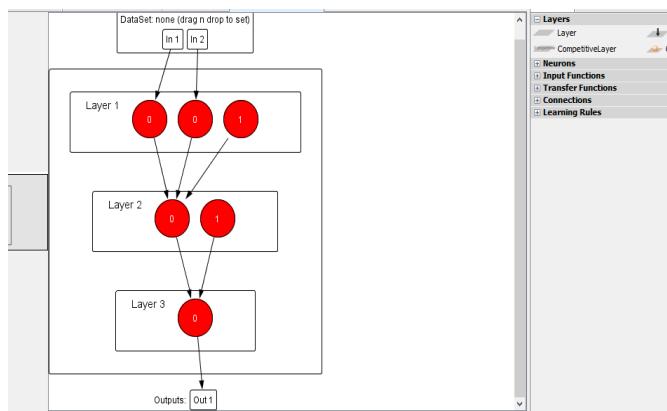


Figure 1. Neural network for generalized function in NeuroPh

We also create a DataSet that will teach and test the neural network. The plurality contains the data presented in Table I.

TABLE I: DATA SET

<b>Input 1</b>	<b>Input 2</b>	<b>Output</b>
-0.5982	0.9870	1.0
-0.2019	0.6210	1.0
0.4715	0.4822	1.0
-0.0982	0.5876	1.0
-0.3566	0.6371	1.0
0.6388	0.4211	1.0
0.6298	0.2815	1.0
-0.4622	0.6166	1.0
-0.0733	0.5582	1.0
-0.5541	0.5125	1.0
-0.4376	0.8781	1.0
-0.2224	0.8885	1.0
0.0935	0.6731	1.0
0.5317	0.5437	1.0
0.4021	0.5164	1.0
0.4756	0.6506	1.0
-0.2338	0.6364	1.0
-0.3158	0.7503	1.0
-0.4735	0.6385	1.0
0.5924	0.8926	1.0
-0.2261	0.7979	1.0
-0.4400	0.5210	1.0
-0.5465	0.7458	1.0
0.4640	0.5107	1.0
-0.1519	0.8122	1.0
0.4854	0.8202	1.0
0.3473	0.7081	1.0
0.4390	0.6282	1.0

-0.2142	0.6436	1.0
0.5738	0.6371	1.0
0.3872	0.5858	1.0
0.3204	0.5353	1.0
-0.2078	0.6513	1.0
-0.1865	0.8175	1.0
0.2475	0.3908	1.0
0.6605	0.8992	1.0
-0.2866	0.7338	1.0
-0.3259	0.3987	1.0
-0.2520	0.6736	1.0
0.3726	0.4979	1.0
-0.2910	1.0437	1.0
-0.3047	0.8686	1.0
-0.2139	1.0932	1.0
-0.3683	0.7564	1.0
-0.4693	0.8878	1.0
0.3935	0.7798	1.0
-0.4564	0.8052	1.0
0.5113	0.7661	1.0
0.2255	0.4645	1.0
0.0146	0.4019	1.0
-0.1917	0.8094	1.0
0.3832	0.7560	1.0
0.4979	0.6133	1.0
0.3534	0.7732	1.0
-0.3472	0.7018	1.0
0.5838	0.7636	1.0
-0.1373	0.7125	1.0
0.3883	0.4498	1.0
-0.5317	0.6193	1.0
-0.1168	0.8785	1.0
0.5434	0.4117	1.0
-0.4540	0.6651	1.0
-0.2191	0.8348	1.0
0.3049	0.9803	1.0
0.6568	0.7577	1.0
0.6142	0.7504	1.0
-0.4581	0.7797	1.0
-0.2162	0.8863	1.0
-0.2602	0.8101	1.0
0.3188	0.8452	1.0
-0.2373	0.8018	1.0
0.5831	0.7771	1.0
0.0284	0.7579	1.0
-0.4184	0.6804	1.0
0.6741	0.6025	1.0
-0.2528	0.7053	1.0
0.5161	0.6209	1.0
0.2039	0.9164	1.0
-0.1721	1.0088	1.0
0.2727	0.2935	1.0
0.0763	0.5622	1.0
-0.3665	0.6483	1.0
0.4429	0.8009	1.0
-0.1998	0.5430	1.0
-0.5408	0.6529	1.0
-0.0706	1.0030	1.0
0.5072	0.3505	1.0
-0.0605	0.6298	1.0
0.2153	0.6026	1.0
0.4681	0.8718	1.0
-0.2989	0.7367	1.0
0.8613	0.4729	1.0
0.7012	0.7457	1.0
-0.1134	0.6007	1.0
0.3123	0.9076	1.0
-0.1217	0.8411	1.0
0.3687	0.3705	1.0
0.5731	0.4095	1.0
-0.2584	0.6719	1.0
0.3094	0.5082	1.0
0.4332	0.7702	1.0

-0.3045	0.5782	1.0
0.4428	0.5802	1.0
-0.1944	0.8988	1.0
-0.0611	0.7418	1.0
0.0762	0.3539	1.0
0.8583	0.9582	1.0
0.3704	0.7234	1.0
0.5148	0.7620	1.0
0.4313	0.5426	1.0
0.4229	0.6524	1.0
0.2982	0.9345	1.0
0.3713	0.7009	1.0
-0.5153	0.7647	1.0
0.3853	0.6553	1.0
-0.3483	0.5053	1.0
0.6851	0.7807	1.0
-0.3653	0.4570	1.0
-0.4090	0.7423	1.0
0.4357	0.4469	1.0
0.2689	0.4456	1.0
-0.4925	1.0144	1.0
0.0762	0.6380	1.0
0.4923	0.4688	1.0
-0.4025	0.7130	1.0
0.0510	0.1609	0.0
-0.7481	0.0890	0.0
-0.7729	0.2632	0.0
0.2184	0.1271	0.0
0.3727	0.4966	0.0
-0.6293	0.6320	0.0
-0.4331	0.1448	0.0
-0.8415	-0.1913	0.0
0.4753	0.2248	0.0
0.3208	0.3272	0.0
0.3206	0.3341	0.0
-0.8908	0.4117	0.0
0.1785	0.4469	0.0
0.3156	0.3885	0.0
0.5578	0.4727	0.0
0.0319	0.0122	0.0
0.2509	0.3072	0.0
0.2357	0.2249	0.0
-0.0724	0.3338	0.0
0.5044	0.0805	0.0
-0.6322	0.4455	0.0
-0.7678	0.2361	0.0
-0.7002	0.2104	0.0
-0.6471	0.1592	0.0
-0.7674	0.0926	0.0
-0.5179	0.0329	0.0
0.1752	0.3453	0.0
-0.6803	0.4761	0.0
0.0160	0.3217	0.0
-0.7148	0.5142	0.0
0.0784	0.3228	0.0
-0.8087	0.4704	0.0
-0.8421	0.0929	0.0
-0.9859	0.4831	0.0
0.2910	0.3428	0.0
0.2432	0.5149	0.0
-0.6010	0.0506	0.0
-1.2465	0.4592	0.0
-0.8277	0.3619	0.0
-0.6212	-0.1091	0.0
-0.7058	0.6591	0.0
0.0672	0.6057	0.0
0.3051	0.4742	0.0
0.6079	0.3936	0.0
-0.7894	0.1759	0.0
-0.5312	0.4265	0.0
0.2520	0.1703	0.0
-0.5788	0.2655	0.0
-0.8318	0.5445	0.0

-0.6986	0.3857	0.0
-0.7364	0.1186	0.0
-0.9350	0.1137	0.0
0.4396	0.4143	0.0
-0.5469	0.2496	0.0
-0.0841	0.3652	0.0
0.3221	0.6909	0.0
0.1076	0.5795	0.0
-0.7186	0.2565	0.0
-0.8788	0.4506	0.0
-0.6985	0.9505	0.0
0.3976	0.1181	0.0
-0.5045	0.5720	0.0
0.2502	0.3978	0.0
0.6171	0.1019	0.0
0.3183	0.0879	0.0
-0.5745	0.1862	0.0
0.0976	0.5518	0.0
0.4845	0.3537	0.0
0.5240	0.4662	0.0
-0.7814	-0.0753	0.0
-0.4970	0.5995	0.0
-0.9698	0.4662	0.0
0.4354	0.1219	0.0
-0.6794	0.3075	0.0
-0.6253	0.0710	0.0
-0.0232	0.4044	0.0
0.2320	0.7107	0.0
0.0938	0.4667	0.0
0.1423	0.1790	0.0
-0.6169	0.2551	0.0
0.2364	0.5154	0.0
0.3891	0.4043	0.0
-0.9518	-0.0377	0.0
0.2409	0.7195	0.0
0.1245	0.4518	0.0
-0.6057	0.2691	0.0
-0.7140	0.3087	0.0
0.3101	0.3468	0.0
0.1802	0.4620	0.0
-0.4266	0.6472	0.0
0.0614	0.3249	0.0
0.0774	0.3218	0.0
0.4281	0.1345	0.0
-0.8025	0.6688	0.0
0.4014	0.4252	0.0
0.3708	0.2641	0.0
-0.8077	0.4149	0.0
0.5016	0.2393	0.0
0.5824	0.2284	0.0
-0.5914	0.3023	0.0
-0.8704	0.2694	0.0
-0.7209	0.1968	0.0
0.2778	0.2179	0.0
0.3324	0.2735	0.0
-0.1409	0.3925	0.0
-0.5976	0.1479	0.0
-0.8558	0.1451	0.0
-0.8891	0.2690	0.0
0.2135	0.4361	0.0
-0.5347	0.5790	0.0
0.3169	0.3971	0.0
-0.6812	0.0421	0.0
-0.9759	0.4596	0.0
0.4146	0.2714	0.0
0.3275	0.3678	0.0
-0.9321	0.0936	0.0
0.5840	0.4715	0.0
-0.4444	0.2301	0.0
0.2911	0.1937	0.0
-0.5108	0.4150	0.0
-0.9660	0.1793	0.0
0.1874	0.2975	0.0

0.1797	0.4518	0.0
-0.7269	0.3573	0.0
-0.5434	0.41011	0.0

So after learning the network and testing, we get the following result for the average error:

Total Mean Square Error: 0.3828746868257739

The complete result of the neural network test is shown in Table. II.

TABLE II: RESULTS OF A NEURAL NETWORK IN NEUROPH

Input 1	Input 2	Output	Desired output	Error
-0.5982	0.987	0.1358	1	-0.8642
-0.2019	0.621	0.1356	1	-0.8644
0.4715	0.4822	0.1353	1	-0.8647
-0.0982	0.5876	0.1356	1	-0.8644
-0.3566	0.6371	0.1357	1	-0.8643
0.6388	0.4211	0.1353	1	-0.8647
0.6298	0.2815	0.1353	1	-0.8647
-0.4622	0.6166	0.1358	1	-0.8642
-0.0733	0.5582	0.1356	1	-0.8644
-0.5541	0.5125	0.1358	1	-0.8642
-0.4376	0.8781	0.1357	1	-0.8643
-0.2224	0.8885	0.1356	1	-0.8644
0.0935	0.6731	0.1355	1	-0.8645
0.5317	0.5437	0.1353	1	-0.8647
0.4021	0.5164	0.1354	1	-0.8646
0.4756	0.6506	0.1353	1	-0.8647
-0.2338	0.6364	0.1356	1	-0.8644
-0.3158	0.7503	0.1357	1	-0.8643
-0.4735	0.6385	0.1358	1	-0.8642
0.5924	0.8926	0.1353	1	-0.8647
-0.2261	0.7979	0.1356	1	-0.8644
-0.44	0.521	0.1357	1	-0.8643
-0.5465	0.7458	0.1358	1	-0.8642
0.464	0.5107	0.1353	1	-0.8647
-0.1519	0.8122	0.1356	1	-0.8644
0.4854	0.8202	0.1353	1	-0.8647
0.3473	0.7081	0.1354	1	-0.8646
0.439	0.6282	0.1353	1	-0.8647
-0.2142	0.6436	0.1356	1	-0.8644
0.5738	0.6371	0.1353	1	-0.8647
0.3872	0.5858	0.1354	1	-0.8646
0.3204	0.5353	0.1354	1	-0.8646
-0.2078	0.6513	0.1356	1	-0.8644
-0.1865	0.8175	0.1356	1	-0.8644
0.2475	0.3908	0.1354	1	-0.8646
0.6605	0.8992	0.1353	1	-0.8647
-0.2866	0.7338	0.1357	1	-0.8643
-0.3259	0.3987	0.1357	1	-0.8643
-0.252	0.6736	0.1357	1	-0.8643
0.3726	0.4979	0.1354	1	-0.8646
-0.291	1.0437	0.1357	1	-0.8643
-0.3047	0.8686	0.1357	1	-0.8643
-0.2139	1.0932	0.1356	1	-0.8644
-0.3683	0.7564	0.1357	1	-0.8643
-0.4693	0.8878	0.1358	1	-0.8642
0.3935	0.7798	0.1354	1	-0.8646
-0.4564	0.8052	0.1357	1	-0.8643
0.5113	0.7661	0.1353	1	-0.8647
0.2255	0.4645	0.1354	1	-0.8646
0.0146	0.4019	0.1355	1	-0.8645
-0.1917	0.8094	0.1356	1	-0.8644
0.3832	0.756	0.1354	1	-0.8646
0.4979	0.6133	0.1353	1	-0.8647

0.3534	0.7732	0.1354	1	-0.8644
-0.3472	0.7018	0.1357	1	-0.8643
0.5838	0.7636	0.1353	1	-0.8647
-0.1373	0.7125	0.1356	1	-0.8644
0.3883	0.4498	0.1354	1	-0.8646
-0.5317	0.6193	0.1358	1	-0.8642
-0.1168	0.8785	0.1356	1	-0.8644
0.5434	0.4117	0.1353	1	-0.8647
-0.454	0.6651	0.1358	1	-0.8642
-0.2191	0.8348	0.1356	1	-0.8644
0.3049	0.9803	0.1354	1	-0.8646
0.6568	0.7577	0.1353	1	-0.8647
0.6142	0.7504	0.1353	1	-0.8647
-0.4581	0.7797	0.1358	1	-0.8642
-0.2162	0.8863	0.1356	1	-0.8644
-0.2602	0.8101	0.1357	1	-0.8643
0.3188	0.8452	0.1354	1	-0.8646
-0.2373	0.8018	0.1356	1	-0.8644
0.5831	0.7771	0.1353	1	-0.8647
0.0284	0.7579	0.1355	1	-0.8645
-0.4184	0.6804	0.1357	1	-0.8643
0.6741	0.6025	0.1353	1	-0.8647
-0.2528	0.7053	0.1357	1	-0.8643
0.5161	0.6209	0.1353	1	-0.8647
0.2039	0.9164	0.1354	1	-0.8646
-0.1721	1.0088	0.1356	1	-0.8644
0.2727	0.2935	0.1354	1	-0.8646
0.0763	0.5622	0.1355	1	-0.8645
-0.3665	0.6483	0.1357	1	-0.8643
0.4429	0.8009	0.1353	1	-0.8647
-0.1998	0.543	0.1356	1	-0.8644
-0.5408	0.6529	0.1358	1	-0.8642
-0.0706	1.003	0.1356	1	-0.8644
0.5072	0.3505	0.1353	1	-0.8647
-0.0605	0.6298	0.1356	1	-0.8644
0.2153	0.6026	0.1354	1	-0.8646
0.4681	0.8718	0.1353	1	-0.8647
-0.2989	0.7367	0.1357	1	-0.8643
0.8613	0.4729	0.1352	1	-0.8648
0.7012	0.7457	0.1352	1	-0.8648
-0.1134	0.6007	0.1356	1	-0.8644
0.3123	0.9076	0.1354	1	-0.8646
-0.1217	0.8411	0.1356	1	-0.8644
0.3687	0.3705	0.1354	1	-0.8646
0.5731	0.4095	0.1353	1	-0.8647
-0.2584	0.6719	0.1357	1	-0.8643
0.3094	0.5082	0.1354	1	-0.8646
0.4332	0.7702	0.1353	1	-0.8647
-0.3045	0.5782	0.1357	1	-0.8643
0.4428	0.5802	0.1353	1	-0.8647
-0.1944	0.8988	0.1356	1	-0.8644
-0.0611	0.7418	0.1356	1	-0.8644
0.0762	0.3539	0.1355	1	-0.8645
0.8583	0.9582	0.1352	1	-0.8648
0.3704	0.7234	0.1354	1	-0.8646
0.5148	0.762	0.1353	1	-0.8647
0.4313	0.5426	0.1353	1	-0.8647
0.4229	0.6524	0.1354	1	-0.8646
0.2982	0.9345	0.1354	1	-0.8646
0.3713	0.7009	0.1354	1	-0.8646
-0.5153	0.7647	0.1358	1	-0.8642
0.3853	0.6553	0.1354	1	-0.8646
-0.3483	0.5053	0.1357	1	-0.8643
0.6851	0.7807	0.1352	1	-0.8648
-0.3653	0.457	0.1357	1	-0.8643
-0.409	0.7423	0.1357	1	-0.8643
0.4357	0.4469	0.1353	1	-0.8647
0.2689	0.4456	0.1354	1	-0.8646
-0.4925	1.0144	0.1358	1	-0.8642
0.0762	0.638	0.1355	1	-0.8645
0.4923	0.4688	0.1353	1	-0.8647
-0.4025	0.713	0.1357	1	-0.8643
0.051	0.1609	0.1355	0	0.1355

-0.7481	0.089	0.1359	0	0.1359
-0.7729	0.2632	0.1359	0	0.1359
0.2184	0.1271	0.1354	0	0.1354
0.3727	0.4966	0.1354	0	0.1354
-0.6293	0.632	0.1358	0	0.1358
-0.4331	0.1448	0.1357	0	0.1357
-0.8415	-0.1913	0.136	0	0.136
0.4753	0.2248	0.1353	0	0.1353
0.3208	0.3272	0.1354	0	0.1354
0.3206	0.3341	0.1354	0	0.1354
-0.8908	0.4117	0.136	0	0.136
0.1785	0.4469	0.1355	0	0.1355
0.3156	0.3885	0.1354	0	0.1354
0.5578	0.4727	0.1353	0	0.1353
0.0319	0.0122	0.1355	0	0.1355
0.2509	0.3072	0.1354	0	0.1354
0.2357	0.2249	0.1354	0	0.1354
-0.0724	0.3338	0.1356	0	0.1356
0.5044	0.0805	0.1353	0	0.1353
-0.6322	0.4455	0.1358	0	0.1358
-0.7678	0.2361	0.1359	0	0.1359
-0.7002	0.2104	0.1359	0	0.1359
-0.6471	0.1592	0.1359	0	0.1359
-0.7674	0.0926	0.1359	0	0.1359
-0.5179	0.0329	0.1358	0	0.1358
0.1752	0.3453	0.1355	0	0.1355
-0.6803	0.4761	0.1359	0	0.1359
0.016	0.3217	0.1355	0	0.1355
-0.7148	0.5142	0.1359	0	0.1359
0.0784	0.3228	0.1355	0	0.1355
-0.8087	0.4704	0.1359	0	0.1359
-0.8421	0.0929	0.136	0	0.136
-0.9859	0.4831	0.136	0	0.136
0.291	0.3428	0.1354	0	0.1354
0.2432	0.5149	0.1354	0	0.1354
-0.601	0.0506	0.1358	0	0.1358
-1.2465	0.4592	0.1362	0	0.1362
-0.8277	0.3619	0.1359	0	0.1359
-0.6212	-0.1091	0.1358	0	0.1358
-0.7058	0.6591	0.1359	0	0.1359
0.0672	0.6057	0.1355	0	0.1355
0.3051	0.4742	0.1354	0	0.1354
0.6079	0.3936	0.1353	0	0.1353
-0.7894	0.1759	0.1359	0	0.1359
-0.5312	0.4265	0.1358	0	0.1358
0.252	0.1703	0.1354	0	0.1354
-0.5788	0.2655	0.1358	0	0.1358
-0.8318	0.5445	0.1359	0	0.1359
-0.6986	0.3857	0.1359	0	0.1359
-0.7364	0.1186	0.1359	0	0.1359
-0.935	0.1137	0.136	0	0.136
0.4396	0.4143	0.1353	0	0.1353
-0.5469	0.2496	0.1358	0	0.1358
-0.0841	0.3652	0.1356	0	0.1356
0.3221	0.6909	0.1354	0	0.1354
0.1076	0.5795	0.1355	0	0.1355
-0.7186	0.2565	0.1359	0	0.1359
-0.8788	0.4506	0.136	0	0.136
-0.6985	0.9505	0.1359	0	0.1359
0.3976	0.1181	0.1354	0	0.1354
-0.5045	0.572	0.1358	0	0.1358
0.2502	0.3978	0.1354	0	0.1354
0.6171	0.1019	0.1353	0	0.1353
0.3183	0.0879	0.1354	0	0.1354
-0.5745	0.1862	0.1358	0	0.1358
0.0976	0.5518	0.1355	0	0.1355
0.4845	0.3537	0.1353	0	0.1353
0.524	0.4662	0.1353	0	0.1353
-0.7814	-0.0753	0.1359	0	0.1359
-0.497	0.5995	0.1358	0	0.1358
-0.9698	0.4662	0.136	0	0.136
0.4354	0.1219	0.1354	0	0.1354
-0.6794	0.3075	0.1359	0	0.1359

-0.6253	0.071	0.1358	0	0.1358
-0.0232	0.4044	0.1355	0	0.1355
0.232	0.7107	0.1354	0	0.1354
0.0938	0.4667	0.1355	0	0.1355
0.1423	0.179	0.1355	0	0.1355
-0.6169	0.2551	0.1358	0	0.1358
0.2364	0.5154	0.1354	0	0.1354
0.3891	0.4043	0.1354	0	0.1354
-0.9518	-0.0377	0.136	0	0.136
0.2409	0.7195	0.1354	0	0.1354
0.1245	0.4518	0.1355	0	0.1355
-0.6057	0.2691	0.1358	0	0.1358
-0.714	0.3087	0.1359	0	0.1359
0.3101	0.3468	0.1354	0	0.1354
0.1802	0.462	0.1355	0	0.1355
-0.4266	0.6472	0.1357	0	0.1357
0.0614	0.3249	0.1355	0	0.1355
0.0774	0.3218	0.1355	0	0.1355
0.4281	0.1345	0.1354	0	0.1354
-0.8025	0.6688	0.1359	0	0.1359
0.4014	0.4252	0.1354	0	0.1354
0.3708	0.2641	0.1354	0	0.1354
-0.8077	0.4149	0.1359	0	0.1359
0.5016	0.2393	0.1353	0	0.1353
0.5824	0.2284	0.1353	0	0.1353
-0.5914	0.3023	0.1358	0	0.1358
-0.8704	0.2694	0.136	0	0.136
-0.7209	0.1968	0.1359	0	0.1359
0.2778	0.2179	0.1354	0	0.1354
0.3324	0.2735	0.1354	0	0.1354
-0.1409	0.3925	0.1356	0	0.1356
-0.5976	0.1479	0.1358	0	0.1358
-0.8558	0.1451	0.136	0	0.136
-0.8891	0.269	0.136	0	0.136
0.2135	0.4361	0.1354	0	0.1354
-0.5347	0.579	0.1358	0	0.1358
0.3169	0.3971	0.1354	0	0.1354
-0.6812	0.0421	0.1359	0	0.1359
-0.9759	0.4596	0.136	0	0.136
0.4146	0.2714	0.1354	0	0.1354
0.3275	0.3678	0.1354	0	0.1354
-0.9321	0.0936	0.136	0	0.136
0.584	0.4715	0.1353	0	0.1353
-0.4444	0.2301	0.1358	0	0.1358
0.2911	0.1937	0.1354	0	0.1354
-0.5108	0.415	0.1358	0	0.1358
-0.966	0.1793	0.136	0	0.136
0.1874	0.2975	0.1355	0	0.1355
0.1797	0.4518	0.1355	0	0.1355
-0.7269	0.3573	0.1359	0	0.1359
-0.5434	0.4101	0.1358	0	0.1358

After the results of the NeuroPh simulator, we have to implement the neural network in logic programming language Visual Prolog.

A class for the neural network is created. It is made up of three files:

Network.cl - this is the class statement.

Network.i - class interface.

Network.pro - performance of the class. [3,7]

In the dialog that implements the feature, it is described with the following Prolog code:

```
onPushButtonClick(_Source) = button::defaultAction:  
-XOR = network::new(),  
XOR:setExamples( [ network::e(-0.5982,0.9870,1.0),  
network::e(-0.2019,0.6210,1.0),  
network::e(0.4715,0.4822,1.0),
```

network::e(-0.0982,0.5876,1.0),  
network::e(-0.3566,0.6371,1.0),  
network::e(0.6388,0.4211,1.0),  
network::e(0.6298,0.2815,1.0),  
network::e(-0.4622,0.6166,1.0),  
network::e(-0.0733,0.5582,1.0),  
network::e(-0.5541,0.5125,1.0),  
network::e(-0.4376,0.8781,1.0),  
network::e(-0.2224,0.8885,1.0),  
network::e(0.0935,0.6731,1.0),  
network::e(0.5317,0.5437,1.0),  
network::e(0.4021,0.5164,1.0),  
network::e(0.4756,0.6506,1.0),  
network::e(-0.2338,0.6364,1.0),  
network::e(-0.3158,0.7503,1.0),  
network::e(-0.4735,0.6385,1.0),  
network::e(0.5924,0.8926,1.0),  
network::e(-0.2261,0.7979,1.0),  
network::e(-0.4400,0.5210,1.0),  
network::e(-0.5465,0.7458,1.0),  
network::e(0.4640,0.5107,1.0),  
network::e(-0.1519,0.8122,1.0),  
network::e(0.4854,0.8202,1.0),  
network::e(0.3473,0.7081,1.0),  
network::e(0.4390,0.6282,1.0),  
network::e(-0.2142,0.6436,1.0),  
network::e(0.5738,0.6371,1.0),  
network::e(0.3872,0.5858,1.0),  
network::e(0.3204,0.5353,1.0),  
network::e(-0.2078,0.6513,1.0),  
network::e(-0.1865,0.8175,1.0),  
network::e(0.2475,0.3908,1.0),  
network::e(0.6605,0.8992,1.0),  
network::e(-0.2866,0.7338,1.0),  
network::e(-0.3259,0.3987,1.0),  
network::e(-0.2520,0.6736,1.0),  
network::e(0.3726,0.4979,1.0),  
network::e(-0.2910,1.0437,1.0),  
network::e(-0.3047,0.8686,1.0),  
network::e(-0.2139,1.0932,1.0),  
network::e(-0.3683,0.7564,1.0),  
network::e(-0.4693,0.8878,1.0),  
network::e(0.3935,0.7798,1.0),  
network::e(-0.4564,0.8052,1.0),  
network::e(0.5113,0.7661,1.0),  
network::e(0.2255,0.4645,1.0),  
network::e(0.0146,0.4019,1.0),  
network::e(-0.1917,0.8094,1.0),  
network::e(0.3832,0.7560,1.0),  
network::e(0.4979,0.6133,1.0),  
network::e(0.3534,0.7732,1.0),  
network::e(-0.3472,0.7018,1.0),  
network::e(0.5838,0.7636,1.0),  
network::e(-0.1373,0.7125,1.0),  
network::e(0.3883,0.4498,1.0),  
network::e(-0.5317,0.6193,1.0),  
network::e(-0.1168,0.8785,1.0),  
network::e(0.5434,0.4117,1.0),  
network::e(-0.4540,0.6651,1.0),  
network::e(-0.2191,0.8348,1.0),  
network::e(0.3049,0.9803,1.0),  
network::e(0.6568,0.7577,1.0),  
network::e(0.6142,0.7504,1.0),  
network::e(-0.4581,0.7797,1.0),  
network::e(-0.2162,0.8863,1.0),  
network::e(-0.2602,0.8101,1.0),  
network::e(0.3188,0.8452,1.0),  
network::e(-0.2373,0.8018,1.0),  
network::e(0.5831,0.7771,1.0),  
network::e(0.0284,0.7579,1.0),  
network::e(-0.4184,0.6804,1.0),  
network::e(0.6741,0.6025,1.0),  
network::e(-0.2528,0.7053,1.0),  
network::e(0.5161,0.6209,1.0),  
network::e(0.2039,0.9164,1.0),  
network::e(-0.1721,1.0088,1.0),  
network::e(0.2727,0.2935,1.0),  
network::e(0.0763,0.5622,1.0),  
network::e(-0.3665,0.6483,1.0),  
network::e(0.4429,0.8009,1.0),  
network::e(-0.1998,0.5430,1.0),  
network::e(-0.5408,0.6529,1.0),  
network::e(-0.0706,1.0030,1.0),  
network::e(0.5072,0.3505,1.0),  
network::e(-0.0605,0.6298,1.0),  
network::e(0.2153,0.6026,1.0),  
network::e(0.4681,0.8718,1.0),  
network::e(-0.2989,0.7367,1.0),  
network::e(0.8613,0.4729,1.0),  
network::e(0.7012,0.7457,1.0),  
network::e(-0.1134,0.6007,1.0),  
network::e(0.3123,0.9076,1.0),  
network::e(-0.1217,0.8411,1.0),  
network::e(0.3687,0.3705,1.0),  
network::e(0.5731,0.4095,1.0),  
network::e(-0.2584,0.6719,1.0),  
network::e(0.3094,0.5082,1.0),  
network::e(0.4332,0.7702,1.0),  
network::e(-0.3045,0.5782,1.0),  
network::e(0.4428,0.5802,1.0),  
network::e(-0.1944,0.8988,1.0),  
network::e(-0.0611,0.7418,1.0),  
network::e(0.0762,0.3539,1.0),  
network::e(0.8583,0.9582,1.0),  
network::e(0.3704,0.7234,1.0),  
network::e(0.5148,0.7620,1.0),  
network::e(0.4313,0.5426,1.0),  
network::e(0.4229,0.6524,1.0),  
network::e(0.2982,0.9345,1.0),  
network::e(0.3713,0.7009,1.0),  
network::e(-0.5153,0.7647,1.0),  
network::e(0.3853,0.6553,1.0),  
network::e(-0.3483,0.5053,1.0),  
network::e(0.6851,0.7807,1.0),  
network::e(-0.3653,0.4570,1.0),  
network::e(-0.4090,0.7423,1.0),  
network::e(0.4357,0.4469,1.0),  
network::e(0.2689,0.4456,1.0),  
network::e(-0.4925,1.0144,1.0),  
network::e(0.0762,0.6380,1.0),  
network::e(0.4923,0.4688,1.0),  
network::e(-0.4025,0.7130,1.0),  
network::e(0.0510,0.1609,0.0),  
network::e(-0.7481,0.0890,0.0),

network::e(-0.7729,0.2632,0.0),  
network::e(0.2184,0.1271,0.0),  
network::e(0.3727,0.4966,0.0),  
network::e(-0.6293,0.6320,0.0),  
network::e(-0.4331,0.1448,0.0),  
network::e(-0.8415,-0.1913,0.0),  
network::e(0.4753,0.2248,0.0),  
network::e(0.3208,0.3272,0.0),  
network::e(0.3206,0.3341,0.0),  
network::e(-0.8908,0.4117,0.0),  
network::e(0.1785,0.4469,0.0),  
network::e(0.3156,0.3885,0.0),  
network::e(0.5578,0.4727,0.0),  
network::e(0.0319,0.0122,0.0),  
network::e(0.2509,0.3072,0.0),  
network::e(0.2357,0.2249,0.0),  
network::e(-0.0724,0.3338,0.0),  
network::e(0.5044,0.0805,0.0),  
network::e(-0.6322,0.4455,0.0),  
network::e(-0.7678,0.2361,0.0),  
network::e(-0.7002,0.2104,0.0),  
network::e(-0.6471,0.1592,0.0),  
network::e(-0.7674,0.0926,0.0),  
network::e(-0.5179,0.0329,0.0),  
network::e(0.1752,0.3453,0.0),  
network::e(-0.6803,0.4761,0.0),  
network::e(0.0160,0.3217,0.0),  
network::e(-0.7148,0.5142,0.0),  
network::e(0.0784,0.3228,0.0),  
network::e(-0.8087,0.4704,0.0),  
network::e(-0.8421,0.0929,0.0),  
network::e(-0.9859,0.4831,0.0),  
network::e(0.2910,0.3428,0.0),  
network::e(0.2432,0.5149,0.0),  
network::e(-0.6010,0.0506,0.0),  
network::e(-1.2465,0.4592,0.0),  
network::e(-0.8277,0.3619,0.0),  
network::e(-0.6212,-0.1091,0.0),  
network::e(-0.7058,0.6591,0.0),  
network::e(0.0672,0.6057,0.0),  
network::e(0.3051,0.4742,0.0),  
network::e(0.6079,0.3936,0.0),  
network::e(-0.7894,0.1759,0.0),  
network::e(-0.5312,0.4265,0.0),  
network::e(0.2520,0.1703,0.0),  
network::e(-0.5788,0.2655,0.0),  
network::e(-0.8318,0.5445,0.0),  
network::e(-0.6986,0.3857,0.0),  
network::e(-0.7364,0.1186,0.0),  
network::e(-0.9350,0.1137,0.0),  
network::e(0.4396,0.4143,0.0),  
network::e(-0.5469,0.2496,0.0),  
network::e(-0.0841,0.3652,0.0),  
network::e(0.3221,0.6909,0.0),  
network::e(0.1076,0.5795,0.0),  
network::e(-0.7186,0.2565,0.0),  
network::e(-0.8788,0.4506,0.0),  
network::e(-0.6985,0.9505,0.0),  
network::e(0.3976,0.1181,0.0),  
network::e(-0.5045,0.5720,0.0),  
network::e(0.2502,0.3978,0.0),  
network::e(0.6171,0.1019,0.0),  
network::e(0.3183,0.0879,0.0),  
network::e(-0.5745,0.1862,0.0),  
network::e(0.0976,0.5518,0.0),  
network::e(0.4845,0.3537,0.0),  
network::e(0.5240,0.4662,0.0),  
network::e(-0.7814,-0.0753,0.0),  
network::e(-0.4970,0.5995,0.0),  
network::e(-0.9698,0.4662,0.0),  
network::e(0.4354,0.1219,0.0),  
network::e(-0.6794,0.3075,0.0),  
network::e(-0.6253,0.0710,0.0),  
network::e(-0.0232,0.4044,0.0),  
network::e(0.2320,0.7107,0.0),  
network::e(0.0938,0.4667,0.0),  
network::e(0.1423,0.1790,0.0),  
network::e(-0.6169,0.2551,0.0),  
network::e(0.2364,0.5154,0.0),  
network::e(0.3891,0.4043,0.0),  
network::e(-0.9518,-0.0377,0.0),  
network::e(0.2409,0.7195,0.0),  
network::e(0.1245,0.4518,0.0),  
network::e(-0.6057,0.2691,0.0),  
network::e(-0.7140,0.3087,0.0),  
network::e(0.3101,0.3468,0.0),  
network::e(0.1802,0.4620,0.0),  
network::e(-0.4266,0.6472,0.0),  
network::e(0.0614,0.3249,0.0),  
network::e(0.0774,0.3218,0.0),  
network::e(0.4281,0.1345,0.0),  
network::e(-0.8025,0.6688,0.0),  
network::e(0.4014,0.4252,0.0),  
network::e(0.3708,0.2641,0.0),  
network::e(-0.8077,0.4149,0.0),  
network::e(0.5016,0.2393,0.0),  
network::e(0.5824,0.2284,0.0),  
network::e(-0.5914,0.3023,0.0),  
network::e(-0.8704,0.2694,0.0),  
network::e(-0.7209,0.1968,0.0),  
network::e(0.2778,0.2179,0.0),  
network::e(0.3324,0.2735,0.0),  
network::e(-0.1409,0.3925,0.0),  
network::e(-0.5976,0.1479,0.0),  
network::e(-0.8558,0.1451,0.0),  
network::e(-0.8891,0.2690,0.0),  
network::e(0.2135,0.4361,0.0),  
network::e(-0.5347,0.5790,0.0),  
network::e(0.3169,0.3971,0.0),  
network::e(-0.6812,0.0421,0.0),  
network::e(-0.9759,0.4596,0.0),  
network::e(0.4146,0.2714,0.0),  
network::e(0.3275,0.3678,0.0),  
network::e(-0.9321,0.0936,0.0),  
network::e(0.5840,0.4715,0.0),  
network::e(-0.4444,0.2301,0.0),  
network::e(0.2911,0.1937,0.0),  
network::e(-0.5108,0.4150,0.0),  
network::e(-0.9660,0.1793,0.0),  
network::e(0.1874,0.2975,0.0),  
network::e(0.1797,0.4518,0.0),  
network::e(-0.7269,0.3573,0.0),  
network::e(-0.5434,0.4101,0.0)],  
training(XOR, 0.001, 0.5),

```

XOR:usit( [-0.5982,0.9870], XOR1), stdio::nl,
stdio::write("xor(-
0.5982,0.9870)=", XOR1), stdio::nl, stdio::nl,
stdio::write("xor(-
XOR:usit( [-0.2019,0.6210], XOR2), stdio::nl,
stdio::write("xor(-
0.2019,0.6210)=", XOR2), stdio::nl, stdio::nl,
stdio::write("xor(0.4715,0.4822)", XOR3), stdio::nl,
stdio::write("xor(0.4715,0.4822)=", XOR3), stdio::nl
, stdio::nl,
stdio::write("xor(-
XOR:usit( [-0.0982,0.5876], XOR4), stdio::nl,
stdio::write("xor(-
0.0982,0.5876)=", XOR4), stdio::nl, stdio::nl,
stdio::write("xor(-
XOR:usit( [-0.3566,0.6371], XOR5), stdio::nl,
stdio::write("xor(-
0.3566,0.6371)=", XOR5), stdio::nl, stdio::nl,
stdio::write("xor(-
XOR:usit( [0.6388,0.4211], XOR6), stdio::nl,
stdio::write("xor(0.6388,0.4211)=", XOR6), stdio::nl
, stdio::nl,
stdio::write("xor(-
XOR:usit( [0.6298,0.2815], XOR7), stdio::nl,
stdio::write("xor(0.6298,0.2815)=", XOR7), stdio::nl
, stdio::nl,
stdio::write("xor(-
0.4622,0.6166)=", XOR8), stdio::nl,
stdio::write("xor(-
0.4622,0.6166)=", XOR8), stdio::nl, stdio::nl,
stdio::write("xor(-
XOR:usit( [-0.0733,0.5582], XOR9), stdio::nl,
stdio::write("xor(-
0.0733,0.5582)=", XOR9), stdio::nl, stdio::nl,
stdio::write("xor(-
0.5541,0.5125)=", XOR10), stdio::nl,
stdio::write("xor(-
0.5541,0.5125)=", XOR10), stdio::nl, stdio::nl,
stdio::write("xor(-
XOR:usit( [-0.4376,0.8781], XOR11), stdio::nl,
stdio::write("xor(-
0.4376,0.8781)=", XOR11), stdio::nl, stdio::nl,
stdio::write("xor(-
XOR:usit( [-0.2224,0.8885], XOR12), stdio::nl,
stdio::write("xor(-
0.2224,0.8885)=", XOR12), stdio::nl, stdio::nl,
stdio::write("xor(-
XOR:usit( [0.0935,0.6731], XOR13), stdio::nl,
stdio::write("xor(0.0935,0.6731)=", XOR13), stdio::
nl, stdio::nl,
stdio::write("xor(-
XOR:usit( [0.5317,0.5437], XOR14), stdio::nl,
stdio::write("xor(0.5317,0.5437)=", XOR14), stdio::
nl, stdio::nl,
stdio::write("xor(-
XOR:usit( [0.4021,0.5164], XOR15), stdio::nl,
stdio::write("xor(0.4021,0.5164)=", XOR15), stdio::
nl, stdio::nl,
stdio::write("xor(-
XOR:usit( [0.4756,0.6506], XOR16), stdio::nl,
stdio::write("xor(0.4756,0.6506)=", XOR16), stdio::
nl, stdio::nl,
stdio::write("xor(-
XOR:usit( [-0.2338,0.6364], XOR17), stdio::nl,
stdio::write("xor(-
0.2338,0.6364)=", XOR17), stdio::nl, stdio::nl,
stdio::write("xor(-
XOR:usit( [-0.3158,0.7503], XOR18), stdio::nl,
stdio::write("xor(-
0.3158,0.7503)=", XOR18), stdio::nl, stdio::nl,
stdio::write("xor(-
XOR:usit( [-0.4735,0.6385], XOR19), stdio::nl,
stdio::write("xor(-
0.4735,0.6385)=", XOR19), stdio::nl, stdio::nl,
stdio::write("xor(0.5924,0.8926)", XOR20), stdio::nl,
stdio::write("xor(0.5924,0.8926)=", XOR20), stdio::
nl, stdio::nl,
stdio::write("xor(-
XOR:usit( [-0.2261,0.7979], XOR21), stdio::nl,
stdio::write("xor(-

```

```

0.2261,0.7979)=", XOR21), stdio::nl, stdio::nl,
stdio::write("xor(-
XOR:usit( [-0.4400,0.5210], XOR22), stdio::nl,
stdio::write("xor(-
0.4400,0.5210)=", XOR22), stdio::nl, stdio::nl,
stdio::write("xor(-
XOR:usit( [-0.5465,0.7458], XOR23), stdio::nl,
stdio::write("xor(-
0.5465,0.7458)=", XOR23), stdio::nl, stdio::nl,
stdio::write("xor(-
XOR:usit( [0.4640,0.5107], XOR24), stdio::nl,
stdio::write("xor(0.4640,0.5107)=", XOR24), stdio::
nl, stdio::nl,
stdio::write("xor(-
XOR:usit( [-0.1519,0.8122], XOR25), stdio::nl,
stdio::write("xor(-
0.1519,0.8122)=", XOR25), stdio::nl, stdio::nl.

```

The result we get after training and testing the neural network is as follows. With 25 positive examples from the training set, we get the result shown in Figure. 2 and Table. III.

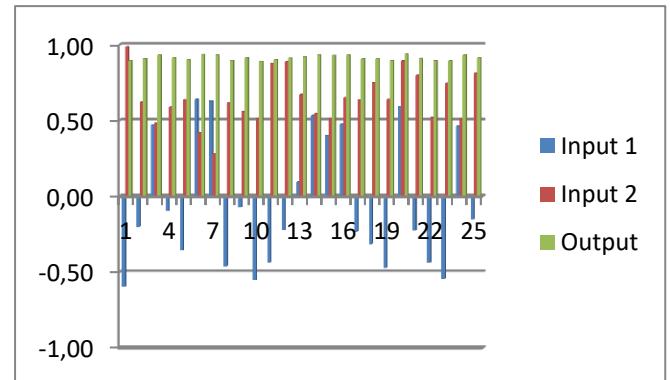


Figure 2. Results of 25 positive input examples

TABLE III. RESULTS OF 25 POSITIVE INPUT EXAMPLES

Input 1	Input 2	Output
-0,60	0,99	0,90
-0,20	0,62	0,91
0,47	0,48	0,93
-0,10	0,59	0,91
-0,36	0,64	0,90
0,64	0,42	0,94
0,63	0,28	0,93
-0,46	0,62	0,90
-0,07	0,56	0,91
-0,55	0,51	0,89
-0,44	0,88	0,90
-0,22	0,89	0,91
0,09	0,67	0,92
0,53	0,54	0,93
0,40	0,52	0,93
0,48	0,65	0,93
-0,23	0,64	0,91
-0,32	0,75	0,91
-0,47	0,64	0,90
0,59	0,89	0,94
-0,23	0,80	0,91
-0,44	0,52	0,90
-0,55	0,75	0,89
0,46	0,51	0,93
-0,15	0,81	0,91

With 30 examples from the training set - 15 positive and 15 negative, we get the result shown in Figure. 3. and in Table IV.

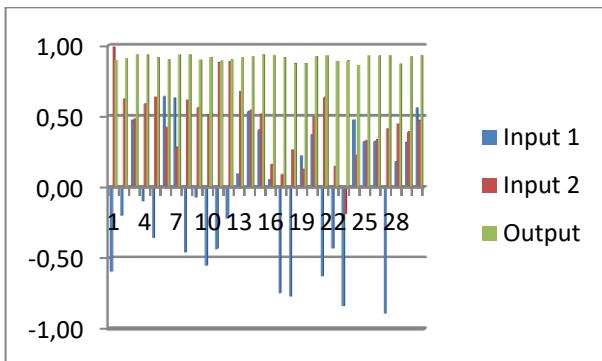


Figure 3. Results of 30 training examples

TABLE IV. RESULTS OF 30 TRAINING EXAMPLES

Input 1	Input 2	Output
-0,60	0,99	0,90
-0,20	0,62	0,91
0,47	0,48	0,93
-0,10	0,59	0,93
-0,36	0,64	0,91
0,64	0,42	0,90
0,63	0,28	0,94
-0,46	0,62	0,93
-0,07	0,56	0,90
-0,55	0,51	0,91
-0,44	0,88	0,89
-0,22	0,89	0,90
0,09	0,67	0,91
0,53	0,54	0,92
0,40	0,52	0,93
0,05	0,16	0,93
-0,75	0,09	0,91
-0,77	0,26	0,87
0,22	0,13	0,87
0,37	0,50	0,92
-0,63	0,63	0,93
-0,43	0,14	0,89
-0,84	-0,19	0,89
0,48	0,22	0,86
0,32	0,33	0,93
0,32	0,33	0,93
-0,89	0,41	0,93
0,18	0,45	0,87
0,32	0,39	0,92
0,56	0,47	0,93

With 4 sample inputs that are not from the training set:

Total Err: 0.0950636151519453

TABLE V. RESULTS FOR 4 INPUT EXAMPLES

Input 1	Input 2	Output
0	0	0.909507490047197
0	1	0.921189000491375
1	0	0.940641576665331
1	1	0.946039530567931

It can be seen that in all three cases the average error is the same. The neural network implemented on

Visual Prolog gives better results than the NeuroPh simulator.

To complete the study, we compare these results with the results of a neural network implemented on the object-oriented Java programming language.

The neural network is implemented with several classes of several .java files. The base class has the following program code:

```
public class Main {
    public static void main(String [] args){
        System.out.println("Starting neural network
sample... ");

        float[][] x
        DataUtils.readInputsFromFile("data/x.txt");
        int[] t
        DataUtils.readOutputsFromFile("data/t.txt");

        NeuralNetwork neuralNetwork = new
        NeuralNetwork(x, t, new INeuralNetworkCallback() {
            @Override
            public void success(Result result) {
                float[] valueToPredict = new float[] {-
                0.205f, 0.780f};
                System.out.println("Success percentage: " +
                result.getSuccessPercentage());
                System.out.println("Predicted result: " +
                result.predictValue(valueToPredict));
            }

            @Override
            public void failure(Error error) {
                System.out.println("Error: " +
                error.getDescription());
            }
        });

        neuralNetwork.startLearning();
    }
}
```

For the training and testing of the neural network, the plurality set forth in Table I is again used. The success rate of this neural network realized on Java is 80%.

A comparison of the mean errors of the three realized neural networks is given in Table VI.

TABLE VI. COMPARISON OF THE AVERAGE ERROR OF NEURAL NETWORKS

Tools	Average error of neural networks
NeuroPh	0.3828746868257739
Visual Prolog	0.0950636151519453
Java	0.8

### III. CONCLUSION

Taking into account the results obtained, we can assert that the highest performance is the neural network implemented in the Visual Prolog logic development

environment. The realized neural network in the logical programming environment produces the smallest average output error.

The neural network implemented on the neural network simulator NeuroPh has a much greater average error. An advantage here is that NeuroPh has a good and lightweight interface as well as a good visual representation of the network. These advantages make the neural network simulator very suitable for the training of beginner programmers on neural networks.

The object - oriented Java programming language on which the third neural network is implemented gives the biggest error of the output. Here the neural network is in several classes - the several files, which also leads to a considerable amount of memory.

Research conducted on the three neural networks tested with the same set of data show that the most appropriate and effective conversion is achieved with the Visual Prolog logic programming language.

#### REFERENCES

- [1] Cubenko G 1989 Approximation by superpositions of a sigmoidal function, Mathematics of control, signals, and systems.  
<http://www.scimagojr.com/journalsearch.php?q=21100376683&tip=sid&clean=0>.
- [2] David J Livingstone 2009 Artificial Neural Network. Methods and Applications
- [3] Eduardo Costa 2009 Visual Prolog 7.2 for Tyros. [http://wiki.visual-prolog.com/Index.php?title=Visual\\_Prolog\\_for\\_Tyros](http://wiki.visual-prolog.com/Index.php?title=Visual_Prolog_for_Tyros)
- [4] Heaton J T 2005 Introduction to Neural Network with Java
- [5] Nachev, A., Hogan, M., Stoyanov, B., Cascade-Correlation Neural Networks for Breast Cancer Diagnosis, Proceedings of the 2011 International Conference on Artificial Intelligence, ICAI 2011, July 18-21, 2011, Las Vegas, USA, 2 Volumes 2011, IC-AI 2011: 475-480. <http://world-comp.org/p2011/ICA4110.pdf>
- [6] Nachev, A., B. Stoyanov, MLP and Default ARTMAP Neural Networks for Business Classification, Intelligent Decision Making System, Proceedings of the 4th International ISKE2009 Conference, World Scientific Series on Computer Engineering and Information Science 2, 179-185. [http://eproceedings.worldscinet.com/9789814295062/9789814295062\\_028.html](http://eproceedings.worldscinet.com/9789814295062/9789814295062_028.html)
- [7] Thomas W.de Boer 2008 A Begginers' Guide to Visual Prolog Groningen [http://wiki.visual-prolog.com/index.php?title=A\\_Beginners\\_Guide\\_to\\_Visual\\_Prolog](http://wiki.visual-prolog.com/index.php?title=A_Beginners_Guide_to_Visual_Prolog)
- [8] Zdravkova, Elitsa., Nenkov, N.V., Comparative analysis of simulators for neuralnetworks Joone and NeuroPh, (On-line: www.cmnt.lv), ISSN 1407-5806, ISSN 1407-5814, TransportandTelecommunication Institute, Riga, Latvia, 2016, 20(1) 16-20. <http://www.scimagojr.com/journalsearch.php?q=21100376683&tip=sid&clean=0>
- [9] Zdravkova, Elitsa., Nenkov, N.V., Research of neuralnetworksimulatorsthroughtwo training data sets, (On-line: www.cmnt.lv), ISSN 1407-5806, ISSN 1407-5814, TransportandTelecommunication Institute, Riga, Latvia, 2016 20(1) 12-15.

#### AUTHORS PROFILE

Elitsa Zdravkova Spasova, 15.02.1988, Bulgaria

Current position, grades: assistant in Shumen University  
University studies: PhD in Informatics  
Scientific interest: Artificial Intelligence, Neural Networks, Genetic Algorithms  
Publications (number or main):7  
Address: Bulgaria, Shumen, Dedeagach 12, ap 24  
Phone: 0885222711

# Using genetic algorithm for shortest path selection with real time traffic flow

<sup>1</sup>Mohammad Aldabbagh, *College of Computer Science and Mathematics, University of Mosul.*

**Abstract**—With the widespread of smart mobile devices and the availability of many applications that provide maps, many programs have spread to find the closest and fastest routes between two points on the map. While the exactness and effectiveness of best path depend on the traffic circumstances, the system needs to add more parameters such as real traffic density and velocity in road. In addition, because of the restricted resources of phone devices, it is not reasonable to be used to calculate the exact optimal solutions by some familiar deterministic algorithms, which are usually used to find the shortest path with a map of reasonable node number. To resolve this issue, this paper put forward to use the genetic algorithm to reduce the computational time. The proposed system use the genetic algorithm to find the shortest path time with miscellaneous situations of real traffic conditions. The genetic algorithm is clearly demonstrate excellent result when applied on many types of map, especially when the number of nodes increased.

## I. INTRODUCTION

Cell phones and tablets became found everywhere as computers by the early 2000s, they also became more advanced to solve the daily problems of their users as calculating taxes, scheduling and aim them to take decisions. However, up to now they cannot compared to normal computer because of their sizes and due to their resources. With this restriction, the execution time for any application became very important. one of the users problems it to find the shortest routes because the (shortest) word mean minimum cost, time and effort.

The shortest path problem can be solve by many algorithms but because of resource limitation of devices, the evolutionary algorithm introduced which they have grown to extremely effective means for resolving optimization problems.

## II. BACKGROUND

### A. GENETIC ALGORITHM

The genetic algorithm is a biologically inspired heuristic search approach to find accurate or approximate solutions [1-9]. The genetic algorithm has on a wide-ranging of applications, including Robotics, finances, Planning and Scheduling [1] [2], pattern recognition [3], Engineering Designs [4] [5], etc. The Steps of the genetic algorithm can be briefly stated as follows:

- Population initialize

- Select the fitness function
- Evaluate each individual in the population with fitness function
- Select the top graded part to breed
- Breed new generation using crossover or mutation
- Replace the worst graded part of population
- Repeat until reach a termination condition.

### Fitness

In the fitness computation step, all individuals of a solution must evaluated on a fitness function. The fitness function measures the quality of the individuals that has generated by Genetic Algorithm.

### Selection

To access the optimal solutions, the best children solutions must selected to be parents in the new population. The selection operation depends on the fitness values in the population.

### Crossover

Crossover is the process of combination of the genetic material of two or more individual solutions. It splits up two individual at  $n$  positions and interchangeably assembles them to a new one [6].

### Mutation

Mutation process based on random changes for individuals [6]. The strength of this disturbance is to keep the solution away of local optima. The mutation process could be controlled by *mutation rate* according to solution spaces.

## B. SHORTEST PATH PROBLEM

The shortest path problem is concern with finding a route between two vertices (nodes) in a graph that have the minimum summation of weights of edges. This problem can be solved simply by Breadth First Search if all edge weights = 1, but here weights can take any value (traffic circumstances). There are many well-known algorithms like Bellman Ford, Dijkstra and Floyd-Warshall Algorithms [7] for solving the problem but when the vertex numbers become too large, the running time will be undesirable.

Chang et al. [8] in 2002 suggested a genetic algorithm to resolve for the shortest path problem. The problem is described as follows:

<sup>1</sup>Software Engineering Department, College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq. (E-mail: m.a.taha@uomosul.edu.iq).

The network can be defined as a  $G = (N, A)$ , where  $N$  is the set of  $n$  nodes (vertices) and  $A$  is the set of edges.  $C = m[C_{ij}]$  is the cost matrix, where  $C_{ij}$  is cost from node  $i$  to node  $j$ . The  $S$  and  $D$  represent the Source and Destination respectively. The  $I_{ij}$  (link indicator) shows whether a route exists between node  $i$  and node  $j$ . If there is a route, then  $I_{ij} = 1$ , otherwise,  $I_{ij} = 0$ .

### III. PROBLEM DESCRIPTION AND SOLUTION

the experiments use SUMO simulator [9] to build a road map grid that include nodes and links which are represent junctions and streets respectively (Figure 1 shows a sample of  $5 \times 5$  node grid and junctions that used in one of experiments designed using SUMO). each junction is controlled by TLS (traffic light system) so we have a grid of traffic lights and that will cause some random obstructions of traffic flow in all streets. All streets (links) are Equipped with induction loop detector(E1) to collect the traffic flow data of the street. All the weights in weights matrix will be updated when the vehicle cross one junction from the chosen path and that will affect the fitness function and taken decision for next node.

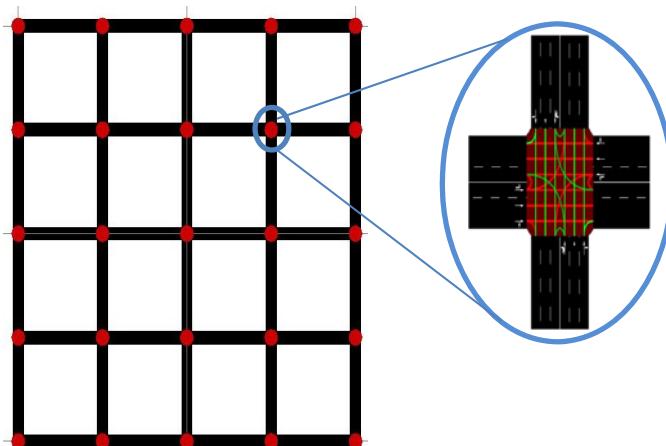


Figure 1: 5x5 Node grid and Junctions

The route from Source Node to Destination node will be represented in chromosome with variable length of chromosome. Figure 2 shows the chromosome structure and that each gene (expect gene 0 for programming needs) represent a node between  $S$  and  $D$ , so all the nodes that have 0 value in  $I$  link indicator matrix will be Excluded from the chromosome.

### The Fitness function

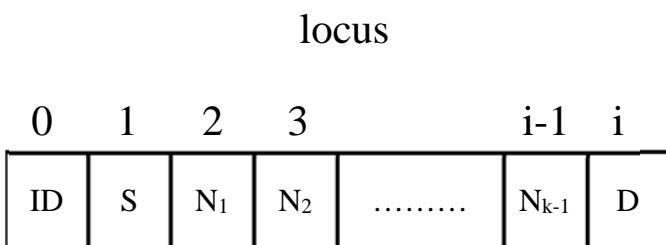


Figure 2: chromosome structure

The fitness function is the quality of the chromosome; therefore, the fitness function is very important for next generation. Ahn and Ramakrishna [8] described the fitness formula as follows:

$$f(i) = \frac{1}{\sum_{j=1}^{L-1} \frac{D(g(i,j), g(i,j+1))}{V(g(i,j), g(i,j+1)) * (1 - Y(g(i,j), g(i,j+1)))}}$$

where

$f(i)$ :  $i$ th chromosome fitness;

$g(i,j)$ :  $i$ th chromosome's  $j$ th gene ;

L: chromosome length;

D: distance between two nodes;

V: velocity limit between two nodes;

Y: Density between two nodes

In all experiments the distance and velocity will be constant and only the density parameter will be variable and will be updated in real time from SUMO detectors.

### Crossover

The algorithm will search for same gene value on two chromosomes in crossover to specify the probable split points, and thus, select one of the points randomly. All the genes after the selected point will be exchanged between chromosome 1 and chromosome 2.

The crossover points are usually different for each parent chromosomes in the crossover phase, in that the crossover may cause a loop that repeated some node as a gene of chromosomes. and those chromosomes will be canceled without affect the crossover rate.

### Mutation

Mutation operation ensures the genetic variety of the population and keep the solution away of local optima. To apply mutation, at each loop (chromosomes) cycle a random number between 0 and 1 is generated and compared to the mutation rate, if less then mutation rate then a mutation point(node) will be selected randomly for the selected chromosome. All the genes before the selected point will be fixed and all genes after the selected point to the S will be generated randomly by initialize function that is used to create the random population.

### Algorithm steps

The algorithm for the working system is as follows:

Begin

S1: Initialize (population size, mutation rate, crossover rate).

S2: Read C values.

S3: Generate the initial population.

S4: Compute fitness for all chromosomes.

S5: Count = 0, G = 1.

S6: loop

S7: Random selection.

S8: Crossover (chromosomes).

S9: Mutation (chromosomes).

S10: Compute fitness for all new chromosomes.

S11: If ( $\min_f[\text{Generation\_now}] = \min_f[\text{Generation\_prev}]$ )  
Counter++; Else stop loop.

S12: Count = 0;  
S13: G++; Jump to S7.  
End

#### IV. THE RESULT

The system had been developed using C# programming language and connected to SUMO using TraCI [10]. Each experiment repeat 50 times and different number of grid implemented (5x5, 10x10, 15x15, 25x25, 50x50) for all the distance and velocity between nodes are constant.

Tables 1 to 5 Show the number of generation needed for each experiment with different crossover and mutation rate.

According to the result, the two parameters (crossover and mutation) are control the time and number of generation required to access the goal because if we increase the mutation or crossover rate that will increase the execution time require to generate the new chromosomes and compare them with fitness function. For most experiment, the best execution times recorded with crossover rate between 20-30% and mutation rate between 1-2%.

grid	Mutation rate %	Crossover rate %	Generation
5x5	1	10	7
		20	6
		30	6
		40	6
		50	5
	2	10	6
		20	6
		30	6
		40	5
		50	5
	3	10	5
		20	5
		30	4
		40	4
		50	3
	4	10	5
		20	4
		30	4
		40	3
		50	2
	5	10	3
		20	3
		30	2
		40	2
		50	2

Table 1: 5x5 grid results

grid	Mutation rate %	Crossover rate %	generation
10x10	1	10	12
		20	12
		30	11
		40	11
		50	11
10x10	2	10	11
		20	11
		30	10
		40	10
		50	8
10x10	3	10	9
		20	9
		30	7
		40	7
		50	6
10x10	4	10	8
		20	7
		30	6
		40	5
		50	5
10x10	5	10	4
		20	4
		30	3
		40	3
		50	3

Table 2: 10x10 grid results

grid	Mutation rate %	Crossover rate %	generation
15x15	1	10	29
		20	27
		30	27
		40	25
		50	23
15x15	2	10	27
		20	27
		30	24
		40	22
		50	20
15x15	3	10	27
		20	26
		30	24
		40	21
		50	20
15x15	4	10	25
		20	25
		30	23

		40	21
		50	19
5	5	10	24
		20	23
		30	21
		40	19
		50	19

Table 3: 15x15 grid results

grid	Mutation rate %	Crossover rate %	generation
25x25	1	10	32
		20	31
		30	31
		40	30
		50	28
	2	10	31
		20	30
		30	29
		40	28
		50	27
	3	10	25
		20	25
		30	23
		40	22
		50	22
	4	10	24
		20	24
		30	23
		40	22
		50	21
	5	10	22
		20	21
		30	20
		40	18
		50	16

Table 4: 25x25 grid results

grid	Mutation rate %	Crossover rate %	generation
50x50	1	10	40
		20	39
		30	39
		40	38
		50	37
	2	10	39
		20	38
		30	37
		40	36
		50	36
	3	10	37
		20	36
		30	35
		40	33

		50	32
4	4	10	33
		20	32
		30	31
		40	30
		50	28
		10	30
5	5	20	29
		30	26
		40	25
		50	24

Table 5: 50x50 grid results

## V. REFERENCES

- [1] P. R. a. D. C. H.-L. Fang, "A promising genetic algorithm approach to job-shop scheduling, rescheduling, & open-shop scheduling," in Proc. 5th Int. Conf. Genetic Algorithms, pp. 375-382, 1993.
- [2] D. A. S. a. G. A. W. G. Mackle, "Application of genetic algorithms to pump scheduling for water supply," in 1st IEE/IEEE Int. Conf. on GA's in Engineering Systems: Innovations and Applications, Sheffield, U.K, pp. 400-405, 1995.
- [3] A. C. L. N. a. K. F. M. S. Kwong, ""Improving local search in genetic algorithms for numerical global optimization using modified GRID-point search technique," 1st IEE/IEEE Int. Conf. on GA's in Engineering Systems: Innovations and Applications, Sheffield, U.K, p. 419423, 1995.
- [4] K. T. a. J. G. J. N. Amaral, "Designing genetic algorithms for the state assignment problem," IEEE Trans. Syst., Man, Cybern., vol. 25, no. 4, 1995.
- [5] V. S. a. o. Vornberger, "Genetic design of VLSI-layouts," in 1st IEE/IEEE Int. Conf. GA's in Engineering Systems: Innovations and Applications, Sheffield, U.K., pp. 430-435. , 1995.
- [6] O. Kramer, Genetic Algorithm Essentials, Springer, 2017.
- [7] V. R. Seth Pettie, "Computing Shortest Paths with Comparisons and Additions," Proceeding SODA '02 Proceedings of the thirteenth annual ACM-SIAM symposium on Discrete algorithms, pp. 267-276 , 2002.
- [8] C. W. A. a. R. Ramakrishnal, "A Genetic Algorithm for Shortest Path Routing Problem and the Sizing of Populations," in Evolutionary Computation, IEEE Transactions, pp. 566-579, 2002.
- [9] J. E. M. B. a. L. B. Daniel Krajzewicz, ""Recent Development and Applications of SUMO - Simulation of Urban MObility", International Journal On Advances in Systems and Measurements, vol. 5 , no. (3&4), pp. 128-138, December 2012.
- [10] M. P. M. R. H. H. S. F. a. J.-P. H. Axel Wegener, "TraCI: A Framework for Coupling Road Traffic and

Network Simulators., "Proceedings of the 11th  
Communications and Networking Simulation  
Symposium, April 2008.

- [11] Y.-W. L. QinQiu Zhang, "An orthogonal genetic algorithm for multimedia multicast routing," in *Evolutionary Computation IEEE Transactions*, pp. 53-62 , 1999 .
- [12] a. U. A. Ismail Rakip Karas, "A genetic algorithm approach for finding the shortest driving time on mobile devices," *Scientific Research and Essays*, vol. 6, no. 2, pp. 394-405, 2011.
- [13] K. A. H. Takagi, "Neural networks in Japan," *Communicat. ACM*, vol. 37, no. 3, pp. 106-112, 1994.
- [14] D. E. G. a. R. E. Smith, "Nonstationary function optimization using genetic dominance and diploidy," in *Proc. 2nd Int. Conf. Genetic Algorithms*, pp. 59-68. , 1987.
- [15] D. E. R. a. M. A. L. B. Widrow, " Neural networks:  
Applications in industry, business, and science," *Communicat. ACM*, vol. 37, no. 3, pp. 93-105, 1994.
- [16] K. S. T. S. K. K. F. Man, "Genetic Algorithms:  
Concepts and Applications," *IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS*, vol. 43, no. 5, pp. 519-534, OCTOBER 1996.

# A software Infrastructure for Multidimensional data Analysis: A Data Modelling Aspect

Prarthana A. Deshkar

*Assistant Professor*

*Yeshwantrao Chavan College of Engineering*

*Nagpur, India*

*prarthana.deshkar@gmail.com*

Dr. Parag S. Deshpande

*Supervisor, CSE dept, G. H. Raisoni College of Engineering, Nagpur, India*

*psdeshpande@cse.vnit.ac.in*

Prof. A. Thomas

*HoD, CSE department*

*G. H. Raisoni College of*

*Engineering, Nagpur, India*

**Abstract** - Rapid changes in the technology lead to increased variety of data sources. These varied data sources generating data in the large volume and with extremely high speed. To accommodate and use this data in decision making systems is the big challenge. To make fullest use of the valuable data generated by different systems, target users of the analysis systems need to be increased. In general knowledge discovery process using the tools which are available requires the handsome expertise in the domain as well as in the technology. The project ITDA (Integrated Tool for Data Analysis) focuses to provide the complete platform for multidimensional data analysis to enhance the decision making process in every domain. This projects provides all the techniques required to perform multidimensional data analysis and avoids the overheads occurred by the traditional cube architecture followed by most of the analytics system. Modelling the available data in the multidimensional form is the basis and crucial step for multidimensional analysis. This work describes the multidimensional modelling aspect and its implementation using ITDA project.

**Keywords** - Multidimensional data analysis, cube, data mining, machine learning, ETL, multidimensional modelling, OLAP.

## I. INTRODUCTION

Due to increased frequency of data generation, data under consideration of analysis is also goes on increasing tremendously. The large size of the data and complexity in data analysis demands an easy platform so that researchers and domain experts can do analysis on their data without the hard core knowledge of information technology. Ad hoc querying or ad hoc reporting is the main need of data analysis. To achieve this data modeling is essential task if the system wants to facilitate the variety of domains. Multidimensional data modeling is the way to provide facility to perform ad hoc analysis. Analyzing multidimensional data is of growing need to extract the knowledge and hence to enable the decision making in various domains. Data analysis process which leads to the enhanced decision making, combine various techniques like statistical techniques, data mining algorithms and machine learning techniques. With all these techniques, presentation of analysis output with attractive visuals is a key part of popular analytics systems. Most of the current multidimensional systems rely on data cubes which are very much resource and time intensive. In this context, ITDA architecture is the solution for multidimensional analysis with the reduced memory and time overheads as compared to the existing systems.

Absorption of high volume of data from variety of sources requires the robust and flexible system. In OLAP terminology the data modelling and data absorption system is called as the Extraction-Transformation-Loading

(ETL) process. The most important bi-product of the ETL process is the metadata. ITDA system uses the on – the – fly architecture for the query generation and hence metadata of the multidimensional model is very crucial component of the system. In a typical analysis environment ETL processes are performed in an ad-hoc, in house fashion or by using some specialized ETL tools. General functionalities of all these tools are identification of relevant information present at the source, extraction of this information, customization and integration of the information coming from multiple sources into a common format, cleansing of the final data set, on the basis of database and business rules, and propagation of the data to relational database which will be used for analysis. In current scenario, organizations might be having number of sources contributing to data collection playing important role in modelling process. The source data might be at different places and it is needed to extract all necessary and data relevant for the analysis. After applying the transformations according to business rule the data is transferred into the target model.

The paper is focusing on this important aspect of any decision making tool, i.e. modeling the data in analysis ready form which may be residing at varied location and may have heterogeneous formats. The organization of this paper is as follows. Firstly in section I we discusses the related work in this area. In section II we give brief introduction of the architecture of the ITDA project along with the basic characteristics of it. In Section III we discuss the conceptual design of the ETL process for ITDA. In next section, section IV we discusses the implementation of the process by considering the case study where data is available in transformed format. And finally we summarize all the contents and discuss the future scope of the system.

## II. RELATED WORK

Multidimensional data analysis system to enhance the efficiency and accuracy of the decision support system is the growing need of today. Many big players of technology like, IBM, Microsoft are having good range of solutions for the same. Every solution is having its own pros and cons. As discussed in [1] most of the multidimensional analysis tools are having stiff curve of learning. Many tools are domain specific. The tools which are having good range of analytical options generally provide the different components for each and every facility which de-motivate the non expert data analysts.

Microstrategy is the leading name in the market of data analysis. Microstrategy provides the component called integrity manager which takes care of the ETL process. It replaces the traditional manual process of data integration. ETL process is the separate component in this tool. Numbers of supporting ETL components are available in the Microstrategy; like, Enterprise Manager ETL, ETL Server, ETL Support, etc. But this may lead to a bit complicated and a costly affair for the research community those are focusing more on analytics and less on technology. [3]

IBM Cognos is also very powerful tool available in the market to perform the multidimensional data analysis. IBM cognos is having different components for each feature like, Cognos for analytics, for business intelligence, for predictive analysis, etc. Cognos Analytics is having a separate data modelling component. This component provides the interface for data extraction from various sources, for transformations and for data validations. [6]

ETL process of ITDA is the integral part of the system to avoid the additional installation and usage overhead of the user.

### III. ITDA SYSTEM ARCHITECTURE

The ITDA system is basically designed to facilitate the researchers and data analyst with the complete package of multidimensional reporting, statistical processing, data mining, machine learning and visualization. This is achieved by the web based system with user friendly and secured environment for the data analyst. This system is functionally independent; it does not require any additional external component or system to complete the task. Also the components of this system are integrated and there is no need to install any of the components separately, which is often common for most of analytics tools.

ITDA system architecture is mainly divided in two parts, data modeling part and data analysis part. Proposed system consists of two main parts containing various components. First is data absorption from different data sources, collection of metadata, and formation of multidimensional model and second is multidimensional analysis on modeled data which further extends to perform statistical analysis and data mining.

Data modeling functionality mainly includes the extraction, transformation and loading (ETL) process. Source data is given to the ETL process and it produces the ready to analyze data. ETL process is responsible to extract the data resides on various sources and in variety of formats. It also performs cleansing and customization of data according to the analysis needs. This process is also responsible to generate the metadata of the ready to analyze data. The proposed system is not going to store the data and the aggregations, hence metadata is having crucial role in this system. Aggregations can be generated on – the – fly by using the metadata.

#### A. ITDA Characteristics

##### *Customized modeling of the data*

Multidimensional modelling of the data according to the business needs is the key of any efficient decision making system. ITDA supports the multiuser system. Each user can model the data in its own way according to business need. In the ITDA terminology the information of the model is conceptualize as the ‘environment’. Single user can have multiple environments for same data so that user gets various views of data for analysis without having complexity of handling number of users for separate business need.

##### *Data absorption options*

ITDA system can accommodate pre-processed data present in flat files where transformation is not required. For such cases it directly loads data in server and collects metadata for that environment. If data is present in multiple sites then this system performs ETL processing during environment creation.

##### *Flexibility in data selection*

Data analyst can have analysis on some particular portion of data by using horizontal partitioning facility given in this system. It allows the user to analyze particular snippet of dataset. It increases performance by reducing number of rows used while running analytical queries or algorithms. User can directly get particular portion of uploaded data by using row filter utility given in the system. This utility allows user to build row filter query without requirement of prior SQL query knowledge. Both these facilities are integrated with the system which can be used after creation of environment.

#### IV. ITDA ETL PROCESS: CONCEPTUAL DESIGN

Process of ETL starts with the understanding of business requirements and the objective of the organization followed by modelling and design of environment for that organization. Modelling and design are defined as representation of key business measurements around its dimensions using dimension modelling. This process decides the level of complexity of transformation based on the source of the data. If data is present at multiple sites then ITDA provides the technique which takes care of extraction of data from multiple sources, transformation and loading.

Last stage of conceptual design part is metadata generation. Metadata contents need to be formulated for a specific multidimensional model. The process decides relationship described by the dimensions, like hierarchical relationship, or sequential relationships. It also gives the level of relationship exists in each component of the dimensional structure. ITDA produces the flat file at the end of the process containing the complete metadata for a multidimensional model created by the user. It also stores the information of the temporal component to create the run time summaries.

##### A. *ETL Algorithm*

During the implementation of the ETL process in ITDA, every correct or missed step is recorded and made available to the user.

- 1) Finalize the ETL processing path
- 2) Finalize the type of data source
- 3) For each any data source map the data source attributes with the dimensional attributes
- 4) Preparation of metadata
- 5) Preparation of configuration file for further processing of model

One of the basic motives behind the ITDA project is to provide the multidimensional analysis platform for non expert data analyst community along with the expert data scientist. The project focuses on interactive, user friendly implementation of the ETL process.

ETL processing path decides whether the data sources are at the same site or located on the different sites. If the data sources are located at different location then the user needs to create one configuration file and based on the instructions given in the file; data will be absorbed. If the data source is at single location then next step is to decide the type of data source like flat file or any other database. Mapping of data source attributes and dimension and fact values are performed and then metadata is generated.

#### V USE OF ETL SERVICE: CASE STUDY FOR FLAT FILE AND DATABASE

ITDA implements the ETL process with highly interactive and user friendly interface. It covers complete ETL process without any programming aspect. Fig 1 gives the main interface of the ITDA system which allows user to initiate the creation of new environment in the system.



Fig 1 ITDA user interface – option to create new environment

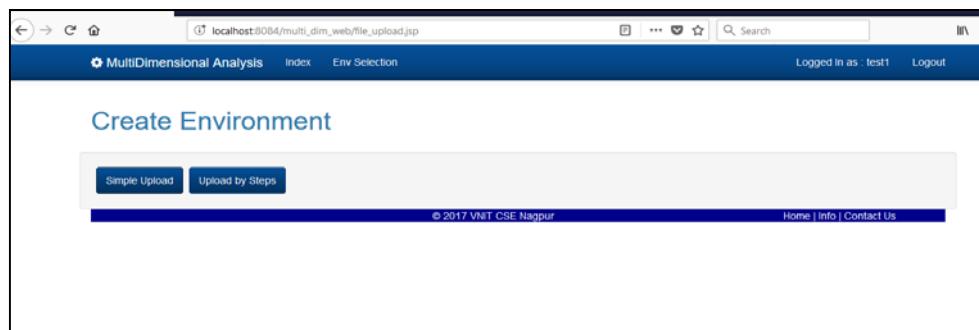


Fig. 2 ITDA interface – selection of ETL processing path

Fig 2 shows the interface which provides the two different paths for ETL process. If the data is available in the already transformed form according to business rule then user will go for the ‘Simple Upload’ option. If the data needs to be extracted from various sources and need the pre-processing according to business rule then the ‘Steps Upload’ option can be the choice.

#### A. Simple upload

This module assumes that data is already in the required form and there is no need of transformation step. For single source data, we can have data in flat files or in database server.

#### B. Flat files

Generally spread sheets or text file formats are used to export data from any database server. If source machine is not accessible from remote location, user can export data in flat files and use those files to create new environment (multidimensional model) in this web tool. ITDA allows user to have data in standard comma separated file or any other flat file with any type of separator. User is allowed to see sample data. Standard query is generated by the system so that user can drop some of the unwanted columns. Successful creation of table enables the metadata collection interface. Figure 3 shows the user interface for uploading csv files to the server. Figure 4 shows the interface with the sample data from the selected file and the standard query generated by the system to extract the file. Analyst can customize the query further.

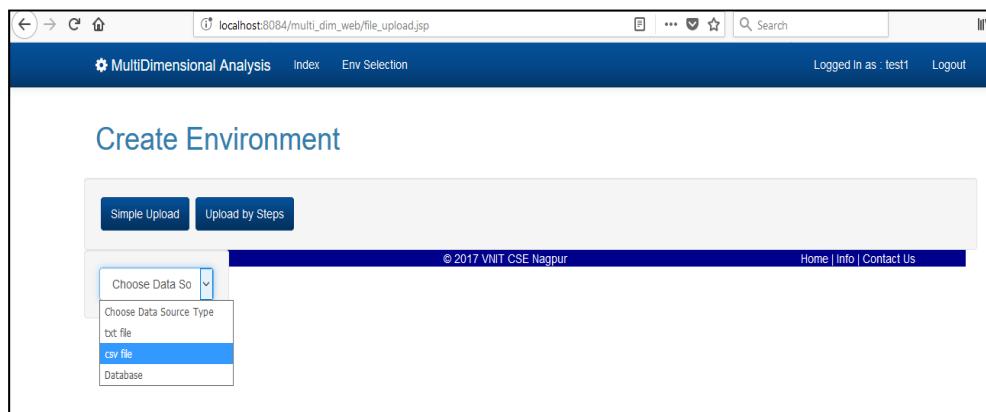


Fig. 3 ITDA interface – selecting file as the data source

Fig. 4 ITDA interface – sample data and editable extraction query

### C. Database

When data source is any database then connection details can be provided so that system can access the data. In this module if source connection and destination connection are same then it will skip data migration process. It avoids extra overload of unnecessary copying entire table. As we are using on the fly architecture, we can use source table for analysis. For analysis we are going to read existing data in OLTP server. In on the fly architecture, we can use same server for OLTP and OLAP processing. This is the biggest advantage of using on the fly architecture. Figure 5 shows database option available in simple upload module.

Fig. 5 ITDA interface – options to map source database for data extraction

#### D. Steps Upload

If we have to take data from multiple sources then we need to have extraction and transformation logic at the server side. In simple upload we were getting pre-processed data so it was much easier to load data in the server and collect metadata. To support collection of data from multiple sites we have steps upload module. This module takes care of extraction, transformation and loading of data at server. In this, we need to upload configuration file to the server containing all details and transformation scripts. We will get all configuration file parameters at the time of conceptual design. Configuration file is in simple text format so that any database user can build it. It is required to keep the process of environment creation as easy as possible. Figure 6 shows steps upload choice



Fig. 6 ITDA interface – option when data needs transformation steps

#### E. Metadata Collection

In order to have multiuser system it is needed to maintain context of every user separately. ITDA ETL process uses specific directory structure for maintaining all the environments created by any user. For every environment there is one flat file for storing the customized operations built by user for performing OLAP. This file is retaining for each environment separately to avoid clash. All the necessary information to operate the models created by a user, a separate directory structure is provided to every user. When user registered for the first time to the system, this complete directory structure is created for that user.

To give metadata user can fill simple html form control to mention dimension names, their hierarchies and time dimension details. Once this data is inserted system can proceed for environment creation.

#### F. ETL for periodic updates

For any ETL system updating data in the server is crucial part. Since OLTP servers will be generating new data continuously. To have analysis on updated data; either the system will change the data available in the server or will add new data keeping earlier data as it is. Here important thing is environment metadata is not changing so metadata collection process can be skipped and directly system can update the data in the required environment.

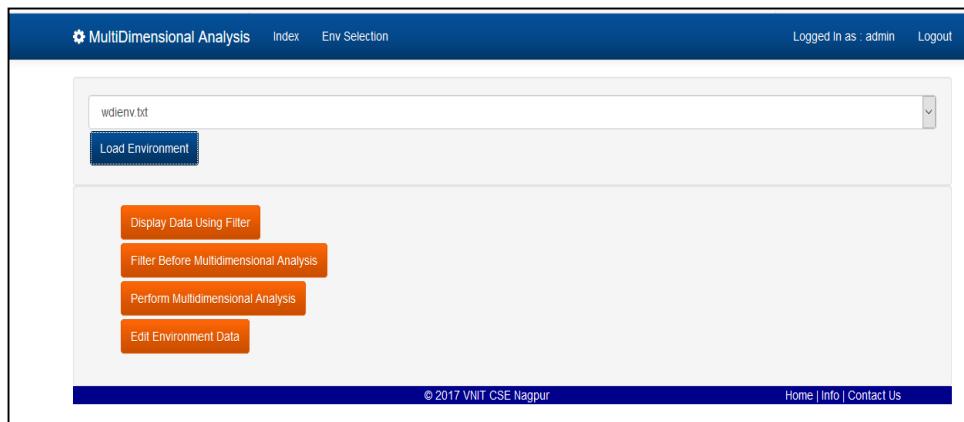


Fig. 7 ITDA interface – options for edit environment

Each user can create any number of environments based on the analysis need. Update process will be invoked for each separate environment. Figure 7 shows environment selection interface and various operations that user can perform after selecting it.

Here this module is for loading new data to the same environment. Figure 8 shows the result after uploading new dataset file to the server. This module flush older data from the table and inserts new data.



Fig. 8 ITDA interface – edit environment option

#### CONCLUSION AND FUTURE WORK

Design of ETL process in ITDA addresses the requirements of efficient extraction, transformation and loading of data from various sources. It also meets with the challenges in assimilating data from heterogeneous data sources, provides an easy to use tool for uploading the existing data set in hand. It successfully collects all metadata parameters required for multidimensional analysis. The designed ETL model can be extended to include facilities of automatic multidimensional modelling where automatic extraction of metadata will be done at the time of load. It can also have context based data generation which collects as well as models the data gathered from web. This data can in turn be tunnelled to multidimensional analysis.

REFERENCES

- [1] Prarthana A. Deshkar, Parag S. Deshpande, A. Thomas, " Multidimensional Data Analysis Facilities and Challenges: A Survey for Data Analysis Tools", International Journal of Computer Applications (0975 – 8887), Volume 179 – No.13, January 2018
- [2] Manasi Vartak, Sajjadur Rahman, Samuel Madden, Aditya Parameswaran, Neoklis Polyzotis "SEEDB: Efficient Data-Driven Visualization Recommendations to Support Visual Analytics", Proceedings of the VLDB Endowment, Vol. 8, No. 13 Copyright 2015 VLDB Endowment 2150-8097/15/09.
- [3] Data Modeling Guide, IBM Cognos Analytics Version 11.0.0, Copyright IBM Corporation 2015, 2017.
- [4] Sandro Fiore, Alessandro D'Anca, Donatello Elia, Cosimo Palazzo, Ian Foster, Dean Williams, Giovanni Aloisio, "Ophidia: a full software stack for scientific data analytics", 978-1-4799-5313-4/14/\$31.00 ©2014 IEEE
- [5] S. Fiore, A. D'Anca, C. Palazzoa,b, I. Fosterc, D. N. Williamsd, G. Aloisioa, "Ophidia: toward big data analytics for eScience", 2013 International Conference on Computational Science, doi: 10.1016/j.procs.2013.05.409, 2013
- [6] Architecture for Enterprise Business Intelligence, an overview of the microstrategy platform architecture for big data, cloud bi, and mobile applications
- [7] Usman AHMED, "Dynamic Cubing for Hierarchical Multidimensional Data Space", PhD thesis, February 2013
- [8] Muntazir Mehdi, Ratnesh Sahay, Wassim Derguech, Edward Curry, "On-The-Fly Generation of Multidimensional Data Cubes for Web of Things", IDEAS '13 October 09 - 11 2013, Barcelona, Spain
- [9] Yang Zhang, Simon Fong, Jinan Fiaidhi, SabahMohammed, "Real-Time Clinical Decision Support Systemwith Data StreamMining", Hindawi Publishing Corporation Journal of Biomedicine and Biotechnology Volume 2012
- [10] Sandra Geisler, Christoph Quix, Stefan Schiffer, Matthias Jarke, "An evaluation framework for traffic information systems based on data streams", 2011 Elsevier Ltd. All rights reserved.
- [11] IBM Cognos Dynamic Cubes, October 2012
- [12] Marta Zorrilla, Diego García-Saiz, "A service oriented architecture to provide data mining services for non-expert data miners",

# Enhanced Generalized Regression Neural Net for Breast Cancer detection

Samira Babaei Ghalejoughi<sup>1</sup>

Department of Electronic Engineering.  
Tabriz Branch, Islamic Azad University.  
Tabriz, Iran.

<sup>1</sup> stu.samira.babaei@iaut.ac.ir

Nasser Lotfivand<sup>\*2</sup>

Department of Electronic Engineering.  
Tabriz Branch, Islamic Azad University.  
Tabriz, Iran.

<sup>\*2</sup> lotfivand@iaut.ac.ir

**Abstract—**In this paper we represent a modified Generalized Regression Artificial Neural net that can recognize all breast cancer of Wisconsin Diagnostic Breast Cancer and Wisconsin Prognostic Breast Cancer correctly. In this method the modified Neural Net trained with 50% of data & 50% for test. But the result is the ability of classify with 100% accuracy. The all 50% train & test data chosen randomly.

This method is based on the fact that calculation in float numbers will remove accuracy. By reducing the number of calculation the accuracy of result increase significantly.

**Keywords-neural network, Generalized regression neural network(GRNN), absolute distance.**

## I. INTRODUCTION (HEADING I)

Pattern classification problems are important application areas of neural networks used as learning systems [1],[2],[3]. Multilayer Perceptrons (MLP), radial basis functions (RBF), probabilistic neural networks (PNN), self-organization maps (SOM), cellular neural networks (CNN), recurrent neural networks and conic section function neural network (CSFNN) are some of these neural networks. In addition to classification problems, function approximation problems are also solved with neural networks. Generalized regression neural network (GRNN) is one of the most popular neural network, used for function approximation. GRNN and PNN are kinds of radial basis function neural networks (RBF-NN) with one pass learning [1]. However they are similar; PNN is used for classification where GRNN is used for continuous function approximation [4]. But in this paper we use GRNN for recognizing.

## II. RELATED WORK

In Back Propagation neural net the neurons trained with gradient descent algorithm the final weight change is (1)

$$\begin{aligned}\Delta W_{pq,k} &= -\eta_{p,q} * (-2 * \alpha * (T_q - \varphi_{q,k}) * \varphi_{q,k} * (1 - \varphi_{q,k}) * \varphi_{p,j}) \\ \Delta W_{hp,j} &= -\eta_{p,q} * \left( \sum_{q=1}^r -2 * \alpha * (T_q - \varphi_{q,k}) * \varphi_{q,k} * \right. \\ &\quad \left. ((1 - \varphi_{q,k}) * \varphi_{p,j} * W_{q,k} * \alpha * \varphi_{p,j} * (1 - \varphi_{p,j}) * x_h) \right)\end{aligned}\quad (1)$$

The  $W_{p,q}$  is the weights between second & third layer &  $W_{h,p}$  is the weights between first & second layer.

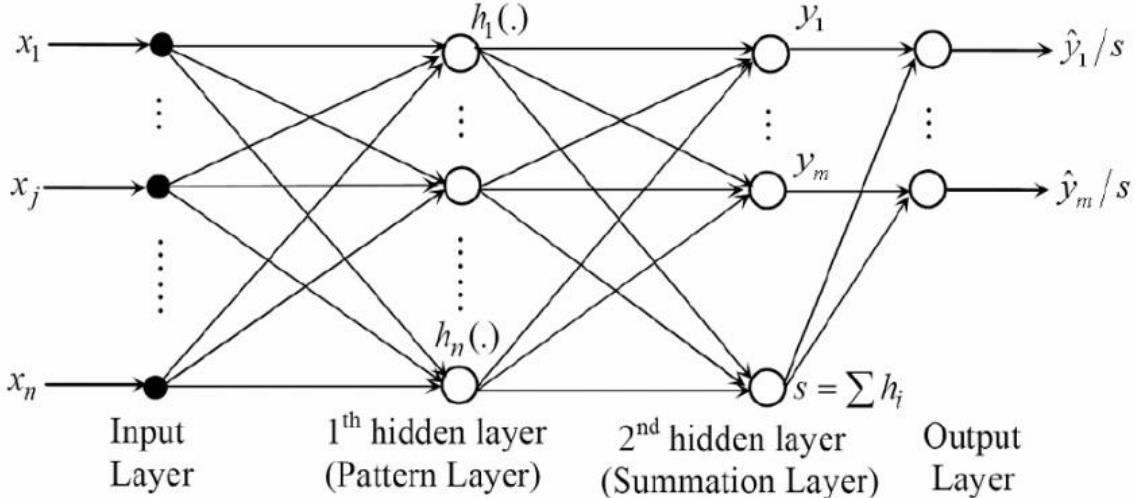
The structure of GRNN Neural net shown in figure(1). The learning algorithm shown in (2).

$$Y(x) = \frac{\sum_{k=1}^N y_k K(x, x_k)}{\sum_{k=1}^N K(x, x_k)} \quad (2)$$

The pros of GRNN is that it can learn in one train. The cons is that it need to save all training data & in some case this need big memory. The  $K(x, x_k)$  is radial base function & the formula for  $K(x, x_k)$  is shown in (3).  $y_k$  is the prediction value for  $x_k$ .  $Y(x)$  is the prediction value for  $x$ .

where  $d_k$  is the squared Euclidean distance between the training samples  $X_k$  and the input  $x$ . In huge data the error increase because of calculating the Euclidean distance.

$$K(x, x_k) = e^{-d_k/2\sigma^2}, \quad d_k = (x - x_k)^T (x - x_k). \quad (3)$$



Figure(1) The GRNN Neural Network

### III. PROPOSED WORK

In the proposed method instead of calculation of Euclidean distance we use absolute distance between samples  $X_k$  & input  $X$ .

### IV. RESULT

To test the simulation we use the Wisconsin Diagnostic Breast Cancer and Wisconsin Prognostic Breast Cancer. First we use 50% of data to train & 50% for test. Then we go one more step & use 40% for train & 60% for test. The result shown in table 1 & table 2.

The result of GRNN enhanced by changing the Euclidean distance to Absolute distance.

TABLE I. RESULT OF SIMULATION FOR 50% TRAIN & 50% TEST.

dataset	WDBC	WPBC	Comment
Number of instances	569	198	
Train Percent	50%	50%	
Test Percent	50%	50%	
Back propagation	95.08%	95.08%	Hidden = 10
Linear SVM	78.12%	79.13%	
Euclidean distance GRNN	94.16%	96.14%	$\sigma = .01$
Absolute distance GRNN	100%	100%	$\sigma = .01$

TABLE II. RESULT OF SIMULATION FOR 40% TRAIN & 60% TEST.

dataset	WDBC	WPBC	Comment
Number of instances	569	198	
Train Percent	40%	40%	
Test Percent	60%	60%	
Back propagation	90.02%	91.49%	Hidden = 10
Linear SVM	78.44%	79.13%	
Euclidean distance GRNN	93.84%	93.84%	$\sigma = .01$
Absolute distance GRNN	100%	100%	$\sigma = .01$

### V. CONCLUSION

In real world we need big data. If we use many calculation the accuracy of computers become low. By reducing the number of calculation we improve the accuracy.

### REFERENCES

- [1] Al-Daoud, E. (2009). A comparison between three neural network models for classification problems. *Journal of Artificial Intelligence*, 2, 56–64.
- [2] Bartlett, P. L. (1998). The sample complexity of pattern classification with neural networks: the size of the weights is more important than the size of the network. *IEEE Transactions on Information Theory*, 44(2), 525–536.
- [3] Specht, D. F. (1990). Probabilistic neural networks. *Neural Networks*, 3, 109–118.
- [4] Specht, D. F. (1991). A general regression neural network. *IEEE Transactions on Neural Networks*, 2(6), 568–576.

## IJCSIS REVIEWERS' LIST

Assist Prof (Dr.) M. Emre Celebi, Louisiana State University in Shreveport, USA  
Dr. Lam Hong Lee, Universiti Tunku Abdul Rahman, Malaysia  
Dr. Shimon K. Modi, Director of Research BSPA Labs, Purdue University, USA  
Dr. Jianguo Ding, Norwegian University of Science and Technology (NTNU), Norway  
Assoc. Prof. N. Jaisankar, VIT University, Vellore, Tamilnadu, India  
Dr. Amogh Kavimandan, The Mathworks Inc., USA  
Dr. Ramasamy Mariappan, Vinayaka Missions University, India  
Dr. Yong Li, School of Electronic and Information Engineering, Beijing Jiaotong University, P.R. China  
Assist. Prof. Sugam Sharma, NIET, India / Iowa State University, USA  
Dr. Jorge A. Ruiz-Vanoye, Universidad Autónoma del Estado de Morelos, Mexico  
Dr. Neeraj Kumar, SMVD University, Katra (J&K), India  
Dr Genge Bela, "Petru Maior" University of Targu Mures, Romania  
Dr. Junjie Peng, Shanghai University, P. R. China  
Dr. Ilhem LENGLIZ, HANA Group - CRISTAL Laboratory, Tunisia  
Prof. Dr. Durgesh Kumar Mishra, Acropolis Institute of Technology and Research, Indore, MP, India  
Dr. Jorge L. Hernández-Ardieta, University Carlos III of Madrid, Spain  
Prof. Dr.C.Suresh Gnana Dhas, Anna University, India  
Dr Li Fang, Nanyang Technological University, Singapore  
Prof. Pijush Biswas, RCC Institute of Information Technology, India  
Dr. Siddhivinayak Kulkarni, University of Ballarat, Ballarat, Victoria, Australia  
Dr. A. Arul Lawrence, Royal College of Engineering & Technology, India  
Dr. Wongyos Keardsri, Chulalongkorn University, Bangkok, Thailand  
Dr. Somesh Kumar Dewangan, CCSVTU Bhilai (C.G.)/ Dimat Raipur, India  
Dr. Hayder N. Jasem, University Putra Malaysia, Malaysia  
Dr. A.V.Senthil Kumar, C. M. S. College of Science and Commerce, India  
Dr. R. S. Karthik, C. M. S. College of Science and Commerce, India  
Dr. P. Vasant, University Technology Petronas, Malaysia  
Dr. Wong Kok Seng, Soongsil University, Seoul, South Korea  
Dr. Praveen Ranjan Srivastava, BITS PILANI, India  
Dr. Kong Sang Kelvin, Leong, The Hong Kong Polytechnic University, Hong Kong  
Dr. Mohd Nazri Ismail, Universiti Kuala Lumpur, Malaysia  
Dr. Rami J. Matarneh, Al-isra Private University, Amman, Jordan  
Dr Ojesanmi Olusegun Ayodeji, Ajayi Crowther University, Oyo, Nigeria  
Dr. Riktesh Srivastava, Skyline University, UAE  
Dr. Oras F. Baker, UCSI University - Kuala Lumpur, Malaysia  
Dr. Ahmed S. Ghiduk, Faculty of Science, Beni-Suef University, Egypt  
and Department of Computer science, Taif University, Saudi Arabia  
Dr. Tirthankar Gayen, IIT Kharagpur, India  
Dr. Huei-Ru Tseng, National Chiao Tung University, Taiwan  
Prof. Ning Xu, Wuhan University of Technology, China  
Dr Mohammed Salem Binwahlan, Hadhramout University of Science and Technology, Yemen  
& Universiti Teknologi Malaysia, Malaysia.  
Dr. Aruna Ranganath, Bhoj Reddy Engineering College for Women, India  
Dr. Hafeezullah Amin, Institute of Information Technology, KUST, Kohat, Pakistan

Prof. Syed S. Rizvi, University of Bridgeport, USA  
Dr. Shahbaz Pervez Chattha, University of Engineering and Technology Taxila, Pakistan  
Dr. Shishir Kumar, Jaypee University of Information Technology, Wakanaghata (HP), India  
Dr. Shahid Mumtaz, Portugal Telecommunication, Instituto de Telecomunicações (IT) , Aveiro, Portugal  
Dr. Rajesh K Shukla, Corporate Institute of Science & Technology Bhopal M P  
Dr. Poonam Garg, Institute of Management Technology, India  
Dr. S. Mehta, Inha University, Korea  
Dr. Dilip Kumar S.M, Bangalore University, Bangalore  
Prof. Malik Sikander Hayat Khiyal, Fatima Jinnah Women University, Rawalpindi, Pakistan  
Dr. Virendra Gomase , Department of Bioinformatics, Padmashree Dr. D.Y. Patil University  
Dr. Irraivan Elamvazuthi, University Technology PETRONAS, Malaysia  
Dr. Saqib Saeed, University of Siegen, Germany  
Dr. Pavan Kumar Gorakavi, IPMA-USA [YC]  
Dr. Ahmed Nabih Zaki Rashed, Menoufia University, Egypt  
Prof. Shishir K. Shandilya, Rukmani Devi Institute of Science & Technology, India  
Dr. J. Komala Lakshmi, SNR Sons College, Computer Science, India  
Dr. Muhammad Sohail, KUST, Pakistan  
Dr. Manjaiah D.H, Mangalore University, India  
Dr. S Santhosh Baboo, D.G.Vaishnav College, Chennai, India  
Prof. Dr. Mokhtar Beldjehem, Sainte-Anne University, Halifax, NS, Canada  
Dr. Deepak Laxmi Narasimha, University of Malaya, Malaysia  
Prof. Dr. Arunkumar Thangavelu, Vellore Institute Of Technology, India  
Dr. M. Azath, Anna University, India  
Dr. Md. Rabiul Islam, Rajshahi University of Engineering & Technology (RUET), Bangladesh  
Dr. Aos Alaa Zaidan Ansaef, Multimedia University, Malaysia  
Dr Suresh Jain, Devi Ahilya University, Indore (MP) India,  
Dr. Mohammed M. Kadhum, Universiti Utara Malaysia  
Dr. Hanumanthappa. J. University of Mysore, India  
Dr. Syed Ishtiaque Ahmed, Bangladesh University of Engineering and Technology (BUET)  
Dr Akinola Solomon Olalekan, University of Ibadan, Ibadan, Nigeria  
Dr. Santosh K. Pandey, The Institute of Chartered Accountants of India  
Dr. P. Vasant, Power Control Optimization, Malaysia  
Dr. Petr Ivankov, Automatika - S, Russian Federation  
Dr. Utkarsh Seetha, Data Infosys Limited, India  
Mrs. Priti Maheshwary, Maulana Azad National Institute of Technology, Bhopal  
Dr. (Mrs) Padmavathi Ganapathi, Avinashilingam University for Women, Coimbatore  
Assist. Prof. A. Neela madheswari, Anna university, India  
Prof. Ganesan Ramachandra Rao, PSG College of Arts and Science, India  
Mr. Kamanashis Biswas, Daffodil International University, Bangladesh  
Dr. Atul Gonsai, Saurashtra University, Gujarat, India  
Mr. Angkoon Phinyomark, Prince of Songkla University, Thailand  
Mrs. G. Nalini Priya, Anna University, Chennai  
Dr. P. Subashini, Avinashilingam University for Women, India  
Assoc. Prof. Vijay Kumar Chakka, Dhirubhai Ambani IICT, Gandhinagar ,Gujarat  
Mr Jitendra Agrawal, : Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal  
Mr. Vishal Goyal, Department of Computer Science, Punjabi University, India  
Dr. R. Baskaran, Department of Computer Science and Engineering, Anna University, Chennai

Assist. Prof. Kanwalvir Singh Dhindsa, B.B.S.B.Engg. College, Fatehgarh Sahib (Punjab), India  
Dr. Jamal Ahmad Dargham, School of Engineering and Information Technology, Universiti Malaysia Sabah  
Mr. Nitin Bhatia, DAV College, India  
Dr. Dhavachelvan Ponnurangam, Pondicherry Central University, India  
Dr. Mohd Faizal Abdollah, University of Technical Malaysia, Malaysia  
Assist. Prof. Sonal Chawla, Panjab University, India  
Dr. Abdul Wahid, AKG Engg. College, Ghaziabad, India  
Mr. Arash Habibi Lashkari, University of Malaya (UM), Malaysia  
Mr. Md. Rajibul Islam, Ibnu Sina Institute, University Technology Malaysia  
Professor Dr. Sabu M. Thampi, .B.S Institute of Technology for Women, Kerala University, India  
Mr. Noor Muhammed Nayeem, Université Lumière Lyon 2, 69007 Lyon, France  
Dr. Himanshu Aggarwal, Department of Computer Engineering, Punjabi University, India  
Prof R. Naidoo, Dept of Mathematics/Center for Advanced Computer Modelling, Durban University of Technology, Durban, South Africa  
Prof. Mydhili K Nair, Visweswaraiah Technological University, Bangalore, India  
M. Prabu, Adhiyamaan College of Engineering/Anna University, India  
Mr. Swakkhar Shatabda, United International University, Bangladesh  
Dr. Abdur Rashid Khan, ICIT, Gomal University, Dera Ismail Khan, Pakistan  
Mr. H. Abdul Shabeer, I-Nautix Technologies, Chennai, India  
Dr. M. Aramudhan, Perunthalaivar Kamarajar Institute of Engineering and Technology, India  
Dr. M. P. Thapliyal, Department of Computer Science, HNB Garhwal University (Central University), India  
Dr. Shahaboddin Shamshirband, Islamic Azad University, Iran  
Mr. Zeashan Hameed Khan, Université de Grenoble, France  
Prof. Anil K Ahlawat, Ajay Kumar Garg Engineering College, Ghaziabad, UP Technical University, Lucknow  
Mr. Longe Olumide Babatope, University Of Ibadan, Nigeria  
Associate Prof. Raman Maini, University College of Engineering, Punjabi University, India  
Dr. Maslin Masrom, University Technology Malaysia, Malaysia  
Sudipta Chattopadhyay, Jadavpur University, Kolkata, India  
Dr. Dang Tuan NGUYEN, University of Information Technology, Vietnam National University - Ho Chi Minh City  
Dr. Mary Lourde R., BITS-PILANI Dubai, UAE  
Dr. Abdul Aziz, University of Central Punjab, Pakistan  
Mr. Karan Singh, Gautam Budtha University, India  
Mr. Avinash Pokhriyal, Uttar Pradesh Technical University, Lucknow, India  
Associate Prof Dr Zuraini Ismail, University Technology Malaysia, Malaysia  
Assistant Prof. Yasser M. Alginahi, Taibah University, Madinah Munawwarah, KSA  
Mr. Dakshina Ranjan Kisku, West Bengal University of Technology, India  
Mr. Raman Kumar, Dr B R Ambedkar National Institute of Technology, Jalandhar, Punjab, India  
Associate Prof. Samir B. Patel, Institute of Technology, Nirma University, India  
Dr. M. Munir Ahmed Rabbani, B. S. Abdur Rahman University, India  
Asst. Prof. Koushik Majumder, West Bengal University of Technology, India  
Dr. Alex Pappachen James, Queensland Micro-nanotechnology center, Griffith University, Australia  
Assistant Prof. S. Hariharan, B.S. Abdur Rahman University, India  
Asst Prof. Jasmine. K. S, R.V. College of Engineering, India  
Mr Naushad Ali Mamode Khan, Ministry of Education and Human Resources, Mauritius  
Prof. Mahesh Goyani, G H Patel College of Engg. & Tech, V.V.N. Anand, Gujarat, India  
Dr. Mana Mohammed, University of Tlemcen, Algeria  
Prof. Jatinder Singh, Universal Institution of Engg. & Tech. CHD, India

Mrs. M. Anandhavalli Gauthaman, Sikkim Manipal Institute of Technology, Majitar, East Sikkim  
Dr. Bin Guo, Institute Telecom SudParis, France  
Mrs. Maleika Mehr Nigar Mohamed Heenaye-Mamode Khan, University of Mauritius  
Prof. Pijush Biswas, RCC Institute of Information Technology, India  
Mr. V. Bala Dhandayuthapani, Mekelle University, Ethiopia  
Dr. Irfan Syamsuddin, State Polytechnic of Ujung Pandang, Indonesia  
Mr. Kavi Kumar Khedo, University of Mauritius, Mauritius  
Mr. Ravi Chandiran, Zagro Singapore Pte Ltd. Singapore  
Mr. Milindkumar V. Sarode, Jawaharlal Darda Institute of Engineering and Technology, India  
Dr. Shamimul Qamar, KSJ Institute of Engineering & Technology, India  
Dr. C. Arun, Anna University, India  
Assist. Prof. M.N.Birje, Basaveshwar Engineering College, India  
Prof. Hamid Reza Naji, Department of Computer Enigneering, Shahid Beheshti University, Tehran, Iran  
Assist. Prof. Debasis Giri, Department of Computer Science and Engineering, Haldia Institute of Technology  
Subhabrata Barman, Haldia Institute of Technology, West Bengal  
Mr. M. I. Lali, COMSATS Institute of Information Technology, Islamabad, Pakistan  
Dr. Feroz Khan, Central Institute of Medicinal and Aromatic Plants, Lucknow, India  
Mr. R. Nagendran, Institute of Technology, Coimbatore, Tamilnadu, India  
Mr. Amnach Khawne, King Mongkut's Institute of Technology Ladkrabang, Ladkrabang, Bangkok, Thailand  
Dr. P. Chakrabarti, Sir Padampat Singhania University, Udaipur, India  
Mr. Nafiz Imtiaz Bin Hamid, Islamic University of Technology (IUT), Bangladesh.  
Shahab-A. Shamshirband, Islamic Azad University, Chalous, Iran  
Prof. B. Priestly Shan, Anna Univeristy, Tamilnadu, India  
Venkatramreddy Velma, Dept. of Bioinformatics, University of Mississippi Medical Center, Jackson MS USA  
Akshi Kumar, Dept. of Computer Engineering, Delhi Technological University, India  
Dr. Umesh Kumar Singh, Vikram University, Ujjain, India  
Mr. Serguei A. Mokhov, Concordia University, Canada  
Mr. Lai Khin Wee, Universiti Teknologi Malaysia, Malaysia  
Dr. Awadhesh Kumar Sharma, Madan Mohan Malviya Engineering College, India  
Mr. Syed R. Rizvi, Analytical Services & Materials, Inc., USA  
Dr. S. Karthik, SNS Collegeof Technology, India  
Mr. Syed Qasim Bukhari, CIMET (Universidad de Granada), Spain  
Mr. A.D.Potgantwar, Pune University, India  
Dr. Himanshu Aggarwal, Punjabi University, India  
Mr. Rajesh Ramachandran, Naipunya Institute of Management and Information Technology, India  
Dr. K.L. Shunmuganathan, R.M.K Engg College , Kavaraipettai ,Chennai  
Dr. Prasant Kumar Pattnaik, KIST, India.  
Dr. Ch. Aswani Kumar, VIT University, India  
Mr. Ijaz Ali Shoukat, King Saud University, Riyadh KSA  
Mr. Arun Kumar, Sir Padam Pat Singhania University, Udaipur, Rajasthan  
Mr. Muhammad Imran Khan, Universiti Teknologi PETRONAS, Malaysia  
Dr. Natarajan Meghanathan, Jackson State University, Jackson, MS, USA  
Mr. Mohd Zaki Bin Mas'ud, Universiti Teknikal Malaysia Melaka (UTeM), Malaysia  
Prof. Dr. R. Geetharamani, Dept. of Computer Science and Eng., Rajalakshmi Engineering College, India  
Dr. Smita Rajpal, Institute of Technology and Management, Gurgaon, India  
Dr. S. Abdul Khader Jilani, University of Tabuk, Tabuk, Saudi Arabia  
Mr. Syed Jamal Haider Zaidi, Bahria University, Pakistan

Dr. N. Devarajan, Government College of Technology, Coimbatore, Tamilnadu, INDIA  
Mr. R. Jagadeesh Kannan, RMK Engineering College, India  
Mr. Deo Prakash, Shri Mata Vaishno Devi University, India  
Mr. Mohammad Abu Naser, Dept. of EEE, IUT, Gazipur, Bangladesh  
Assist. Prof. Prasun Ghosal, Bengal Engineering and Science University, India  
Mr. Md. Golam Kaosar, School of Engineering and Science, Victoria University, Melbourne City, Australia  
Mr. R. Mohammad Shafi, Madanapalle Institute of Technology & Science, India  
Dr. F.Sagayaraj Francis, Pondicherry Engineering College, India  
Dr. Ajay Goel, HIET, Kaithal, India  
Mr. Nayak Sunil Kashibarao, Bahirji Smarak Mahavidyalaya, India  
Mr. Suhas J Manangi, Microsoft India  
Dr. Kalyankar N. V., Yeshwant Mahavidyalaya, Nanded, India  
Dr. K.D. Verma, S.V. College of Post graduate studies & Research, India  
Dr. Amjad Rehman, University Technology Malaysia, Malaysia  
Mr. Rachit Garg, L K College, Jalandhar, Punjab  
Mr. J. William, M.A.M college of Engineering, Trichy, Tamilnadu, India  
Prof. Jue-Sam Chou, Nanhua University, College of Science and Technology, Taiwan  
Dr. Thorat S.B., Institute of Technology and Management, India  
Mr. Ajay Prasad, Sir Padampat Singhania University, Udaipur, India  
Dr. Kamaljit I. Lakhtaria, Atmiya Institute of Technology & Science, India  
Mr. Syed Raifiul Hussain, Ahsanullah University of Science and Technology, Bangladesh  
Mrs Fazeela Tunnis, Najran University, Kingdom of Saudi Arabia  
Mrs Kavita Taneja, Maharishi Markandeshwar University, Haryana, India  
Mr. Maniyar Shiraz Ahmed, Najran University, Najran, KSA  
Mr. Anand Kumar, AMC Engineering College, Bangalore  
Dr. Rakesh Chandra Gangwar, Beant College of Engg. & Tech., Gurdaspur (Punjab) India  
Dr. V V Rama Prasad, Sree Vidyanikethan Engineering College, India  
Assist. Prof. Neetesh Kumar Gupta, Technocrats Institute of Technology, Bhopal (M.P.), India  
Mr. Ashish Seth, Uttar Pradesh Technical University, Lucknow, UP India  
Dr. V V S S Balaram, Sreenidhi Institute of Science and Technology, India  
Mr Rahul Bhatia, Lingaya's Institute of Management and Technology, India  
Prof. Niranjana Reddy. P, KITS, Warangal, India  
Prof. Rakesh. Lingappa, Vijetha Institute of Technology, Bangalore, India  
Dr. Mohammed Ali Hussain, Nimra College of Engineering & Technology, Vijayawada, A.P., India  
Dr. A.Srinivasan, MNM Jain Engineering College, Rajiv Gandhi Salai, Thorapakkam, Chennai  
Mr. Rakesh Kumar, M.M. University, Mullana, Ambala, India  
Dr. Lena Khaled, Zarqa Private University, Aman, Jordan  
Ms. Supriya Kapoor, Patni/Lingaya's Institute of Management and Tech., India  
Dr. Tossapon Boongoen, Aberystwyth University, UK  
Dr. Bilal Alatas, Firat University, Turkey  
Assist. Prof. Jyoti Praaksh Singh, Academy of Technology, India  
Dr. Ritu Soni, GNG College, India  
Dr. Mahendra Kumar, Sagar Institute of Research & Technology, Bhopal, India.  
Dr. Binod Kumar, Lakshmi Narayan College of Tech.(LNCT)Bhopal India  
Dr. Muzhir Shaban Al-Ani, Amman Arab University Amman – Jordan  
Dr. T.C. Manjunath, ATRIA Institute of Tech, India  
Mr. Muhammad Zakarya, COMSATS Institute of Information Technology (CIIT), Pakistan

Assist. Prof. Harmunish Taneja, M. M. University, India  
Dr. Chitra Dhawale , SICSR, Model Colony, Pune, India  
Mrs Sankari Muthukaruppan, Nehru Institute of Engineering and Technology, Anna University, India  
Mr. Aaqif Afzaal Abbasi, National University Of Sciences And Technology, Islamabad  
Prof. Ashutosh Kumar Dubey, Trinity Institute of Technology and Research Bhopal, India  
Mr. G. Appasami, Dr. Pauls Engineering College, India  
Mr. M Yasin, National University of Science and Tech, karachi (NUST), Pakistan  
Mr. Yaser Miaji, University Utara Malaysia, Malaysia  
Mr. Shah Ahsanul Haque, International Islamic University Chittagong (IIUC), Bangladesh  
Prof. (Dr) Syed Abdul Sattar, Royal Institute of Technology & Science, India  
Dr. S. Sasikumar, Roever Engineering College  
Assist. Prof. Monit Kapoor, Maharishi Markandeshwar University, India  
Mr. Nwaocha Vivian O, National Open University of Nigeria  
Dr. M. S. Vijaya, GR Govindarajulu School of Applied Computer Technology, India  
Assist. Prof. Chakresh Kumar, Manav Rachna International University, India  
Mr. Kunal Chadha , R&D Software Engineer, Gemalto, Singapore  
Mr. Mueen Uddin, Universiti Teknologi Malaysia, UTM , Malaysia  
Dr. Dhuha Basheer abdullah, Mosul university, Iraq  
Mr. S. Audithan, Annamalai University, India  
Prof. Vijay K Chaudhari, Technocrats Institute of Technology , India  
Associate Prof. Mohd Ilyas Khan, Technocrats Institute of Technology , India  
Dr. Vu Thanh Nguyen, University of Information Technology, HoChiMinh City, VietNam  
Assist. Prof. Anand Sharma, MITS, Lakshmangarh, Sikar, Rajasthan, India  
Prof. T V Narayana Rao, HITAM Engineering college, Hyderabad  
Mr. Deepak Gour, Sir Padampat Singhania University, India  
Assist. Prof. Amutharaj Joyson, Kalasalingam University, India  
Mr. Ali Balador, Islamic Azad University, Iran  
Mr. Mohit Jain, Maharaja Surajmal Institute of Technology, India  
Mr. Dilip Kumar Sharma, GLA Institute of Technology & Management, India  
Dr. Debojyoti Mitra, Sir padampat Singhania University, India  
Dr. Ali Dehghantanha, Asia-Pacific University College of Technology and Innovation, Malaysia  
Mr. Zhao Zhang, City University of Hong Kong, China  
Prof. S.P. Setty, A.U. College of Engineering, India  
Prof. Patel Rakeshkumar Kantilal, Sankalchand Patel College of Engineering, India  
Mr. Biswajit Bhowmik, Bengal College of Engineering & Technology, India  
Mr. Manoj Gupta, Apex Institute of Engineering & Technology, India  
Assist. Prof. Ajay Sharma, Raj Kumar Goel Institute Of Technology, India  
Assist. Prof. Ramveer Singh, Raj Kumar Goel Institute of Technology, India  
Dr. Hanan Elazhary, Electronics Research Institute, Egypt  
Dr. Hosam I. Faiq, USM, Malaysia  
Prof. Dipti D. Patil, MAEER's MIT College of Engg. & Tech, Pune, India  
Assist. Prof. Devendra Chack, BCT Kumaon engineering College Dwarahat Almora, India  
Prof. Manpreet Singh, M. M. Engg. College, M. M. University, India  
Assist. Prof. M. Sadiq ali Khan, University of Karachi, Pakistan  
Mr. Prasad S. Halgaonkar, MIT - College of Engineering, Pune, India  
Dr. Imran Ghani, Universiti Teknologi Malaysia, Malaysia  
Prof. Varun Kumar Kakar, Kumaon Engineering College, Dwarahat, India

Assist. Prof. Nisheeth Joshi, Apaji Institute, Banasthali University, Rajasthan, India  
Associate Prof. Kunwar S. Vaisla, VCT Kumaon Engineering College, India  
Prof Anupam Choudhary, Bhilai School Of Engg.,Bhilai (C.G.),India  
Mr. Divya Prakash Shrivastava, Al Jabal Al garbi University, Zawya, Libya  
Associate Prof. Dr. V. Radha, Avinashilingam Deemed university for women, Coimbatore.  
Dr. Kasarapu Ramani, JNT University, Anantapur, India  
Dr. Anuraag Awasthi, Jayoti Vidyapeeth Womens University, India  
Dr. C G Ravichandran, R V S College of Engineering and Technology, India  
Dr. Mohamed A. Deriche, King Fahd University of Petroleum and Minerals, Saudi Arabia  
Mr. Abbas Karimi, Universiti Putra Malaysia, Malaysia  
Mr. Amit Kumar, Jaypee University of Engg. and Tech., India  
Dr. Nikolai Stoianov, Defense Institute, Bulgaria  
Assist. Prof. S. Ranichandra, KSR College of Arts and Science, Tiruchencode  
Mr. T.K.P. Rajagopal, Diamond Horse International Pvt Ltd, India  
Dr. Md. Ekramul Hamid, Rajshahi University, Bangladesh  
Mr. Hemanta Kumar Kalita , TATA Consultancy Services (TCS), India  
Dr. Messaouda Azzouzi, Ziane Achour University of Djelfa, Algeria  
Prof. (Dr.) Juan Jose Martinez Castillo, "Gran Mariscal de Ayacucho" University and Acantelys research Group, Venezuela  
Dr. Jatinderkumar R. Saini, Narmada College of Computer Application, India  
Dr. Babak Bashari Rad, University Technology of Malaysia, Malaysia  
Dr. Nighat Mir, Effat University, Saudi Arabia  
Prof. (Dr.) G.M.Nasira, Sasurie College of Engineering, India  
Mr. Varun Mittal, Gemalto Pte Ltd, Singapore  
Assist. Prof. Mrs P. Banumathi, Kathir College Of Engineering, Coimbatore  
Assist. Prof. Quan Yuan, University of Wisconsin-Stevens Point, US  
Dr. Pranam Paul, Narula Institute of Technology, Agarpara, West Bengal, India  
Assist. Prof. J. Ramkumar, V.L.B Janakiammal college of Arts & Science, India  
Mr. P. Sivakumar, Anna university, Chennai, India  
Mr. Md. Humayun Kabir Biswas, King Khalid University, Kingdom of Saudi Arabia  
Mr. Mayank Singh, J.P. Institute of Engg & Technology, Meerut, India  
HJ. Kamaruzaman Jusoff, Universiti Putra Malaysia  
Mr. Nikhil Patrick Lobo, CADES, India  
Dr. Amit Wason, Rayat-Bahra Institute of Engineering & Boi-Technology, India  
Dr. Rajesh Shrivastava, Govt. Benazir Science & Commerce College, Bhopal, India  
Assist. Prof. Vishal Bharti, DCE, Gurgaon  
Mrs. Sunita Bansal, Birla Institute of Technology & Science, India  
Dr. R. Sudhakar, Dr.Mahalingam college of Engineering and Technology, India  
Dr. Amit Kumar Garg, Shri Mata Vaishno Devi University, Katra(J&K), India  
Assist. Prof. Raj Gaurang Tiwari, AZAD Institute of Engineering and Technology, India  
Mr. Hamed Taherdoost, Tehran, Iran  
Mr. Amin Daneshmand Malayeri, YRC, IAU, Malayer Branch, Iran  
Mr. Shantanu Pal, University of Calcutta, India  
Dr. Terry H. Walcott, E-Promag Consultancy Group, United Kingdom  
Dr. Ezekiel U OKIKE, University of Ibadan, Nigeria  
Mr. P. Mahalingam, Caledonian College of Engineering, Oman  
Dr. Mahmoud M. A. Abd Ellatif, Mansoura University, Egypt

Prof. Kunwar S. Vaisla, BCT Kumaon Engineering College, India  
Prof. Mahesh H. Panchal, Kalol Institute of Technology & Research Centre, India  
Mr. Muhammad Asad, Technical University of Munich, Germany  
Mr. AliReza Shams Shafiqh, Azad Islamic university, Iran  
Prof. S. V. Nagaraj, RMK Engineering College, India  
Mr. Ashikali M Hasan, Senior Researcher, CelNet security, India  
Dr. Adnan Shahid Khan, University Technology Malaysia, Malaysia  
Mr. Prakash Gajanan Burade, Nagpur University/ITM college of engg, Nagpur, India  
Dr. Jagdish B.Helonde, Nagpur University/ITM college of engg, Nagpur, India  
Professor, Doctor BOUHORMA Mohammed, Univertsity Abdelmalek Essaadi, Morocco  
Mr. K. Thirumalaivasan, Pondicherry Engg. College, India  
Mr. Umbarkar Anantkumar Janardan, Walchand College of Engineering, India  
Mr. Ashish Chaurasia, Gyan Ganga Institute of Technology & Sciences, India  
Mr. Sunil Taneja, Kurukshetra University, India  
Mr. Fauzi Adi Rafrastra, Dian Nuswantoro University, Indonesia  
Dr. Yaduvir Singh, Thapar University, India  
Dr. Ioannis V. Koskosas, University of Western Macedonia, Greece  
Dr. Vasantha Kalyani David, Avinashilingam University for women, Coimbatore  
Dr. Ahmed Mansour Manasrah, Universiti Sains Malaysia, Malaysia  
Miss. Nazanin Sadat Kazazi, University Technology Malaysia, Malaysia  
Mr. Saeed Rasouli Heikalabad, Islamic Azad University - Tabriz Branch, Iran  
Assoc. Prof. Dhirendra Mishra, SVKM's NMIMS University, India  
Prof. Shapor Zarei, UAE Inventors Association, UAE  
Prof. B.Raja Sarath Kumar, Lenora College of Engineering, India  
Dr. Bashir Alam, Jamia millia Islamia, Delhi, India  
Prof. Anant J Umbarkar, Walchand College of Engg., India  
Assist. Prof. B. Bharathi, Sathyabama University, India  
Dr. Fokrul Alom Mazarbhuiya, King Khalid University, Saudi Arabia  
Prof. T.S.Jeyali Laseeth, Anna University of Technology, Tirunelveli, India  
Dr. M. Balraju, Jawahar Lal Nehru Technological University Hyderabad, India  
Dr. Vijayalakshmi M. N., R.V.College of Engineering, Bangalore  
Prof. Walid Moudani, Lebanese University, Lebanon  
Dr. Saurabh Pal, VBS Purvanchal University, Jaunpur, India  
Associate Prof. Suneet Chaudhary, Dehradun Institute of Technology, India  
Associate Prof. Dr. Manuj Darbari, BBD University, India  
Ms. Prema Selvaraj, K.S.R College of Arts and Science, India  
Assist. Prof. Ms.S.Sasikala, KSR College of Arts & Science, India  
Mr. Sukhvinder Singh Deora, NC Institute of Computer Sciences, India  
Dr. Abhay Bansal, Amity School of Engineering & Technology, India  
Ms. Sumita Mishra, Amity School of Engineering and Technology, India  
Professor S. Viswanadha Raju, JNT University Hyderabad, India  
Mr. Asghar Shahrzad Khashandarag, Islamic Azad University Tabriz Branch, India  
Mr. Manoj Sharma, Panipat Institute of Engg. & Technology, India  
Mr. Shakeel Ahmed, King Faisal University, Saudi Arabia  
Dr. Mohamed Ali Mahjoub, Institute of Engineer of Monastir, Tunisia  
Mr. Adri Jovin J.J., SriGuru Institute of Technology, India  
Dr. Sukumar Senthilkumar, Universiti Sains Malaysia, Malaysia

Mr. Rakesh Bharati, Dehradun Institute of Technology Dehradun, India  
Mr. Shervan Fekri Ershad, Shiraz International University, Iran  
Mr. Md. Safiqul Islam, Daffodil International University, Bangladesh  
Mr. Mahmudul Hasan, Daffodil International University, Bangladesh  
Prof. Mandakini Tayade, UIT, RGTU, Bhopal, India  
Ms. Sarla More, UIT, RGTU, Bhopal, India  
Mr. Tushar Hrishikesh Jaware, R.C. Patel Institute of Technology, Shirpur, India  
Ms. C. Divya, Dr G R Damodaran College of Science, Coimbatore, India  
Mr. Fahimuddin Shaik, Annamacharya Institute of Technology & Sciences, India  
Dr. M. N. Giri Prasad, JNTUCE,Pulivendula, A.P., India  
Assist. Prof. Chintan M Bhatt, Charotar University of Science And Technology, India  
Prof. Sahista Machchhar, Marwadi Education Foundation's Group of institutions, India  
Assist. Prof. Navnish Goel, S. D. College Of Enginnering & Technology, India  
Mr. Khaja Kamaluddin, Sirt University, Sirt, Libya  
Mr. Mohammad Zaidul Karim, Daffodil International, Bangladesh  
Mr. M. Vijayakumar, KSR College of Engineering, Tiruchengode, India  
Mr. S. A. Ahsan Rajon, Khulna University, Bangladesh  
Dr. Muhammad Mohsin Nazir, LCW University Lahore, Pakistan  
Mr. Mohammad Asadul Hoque, University of Alabama, USA  
Mr. P.V.Sarathchand, Indur Institute of Engineering and Technology, India  
Mr. Durgesh Samadhiya, Chung Hua University, Taiwan  
Dr Venu Kuthadi, University of Johannesburg, Johannesburg, RSA  
Dr. (Er) Jasvir Singh, Guru Nanak Dev University, Amritsar, Punjab, India  
Mr. Jasmin Cosic, Min. of the Interior of Una-sana canton, B&H, Bosnia and Herzegovina  
Dr S. Rajalakshmi, Botho College, South Africa  
Dr. Mohamed Sarrab, De Montfort University, UK  
Mr. Basappa B. Kodada, Canara Engineering College, India  
Assist. Prof. K. Ramana, Annamacharya Institute of Technology and Sciences, India  
Dr. Ashu Gupta, Apeejay Institute of Management, Jalandhar, India  
Assist. Prof. Shaik Rasool, Shadan College of Engineering & Technology, India  
Assist. Prof. K. Suresh, Annamacharya Institute of Tech & Sci. Rajampet, AP, India  
Dr . G. Singaravel, K.S.R. College of Engineering, India  
Dr B. G. Geetha, K.S.R. College of Engineering, India  
Assist. Prof. Kavita Choudhary, ITM University, Gurgaon  
Dr. Mehrdad Jalali, Azad University, Mashhad, Iran  
Megha Goel, Shamli Institute of Engineering and Technology, Shamli, India  
Mr. Chi-Hua Chen, Institute of Information Management, National Chiao-Tung University, Taiwan (R.O.C.)  
Assoc. Prof. A. Rajendran, RVS College of Engineering and Technology, India  
Assist. Prof. S. Jaganathan, RVS College of Engineering and Technology, India  
Assoc. Prof. (Dr.) A S N Chakravarthy, JNTUK University College of Engineering Vizianagaram (State University)  
Assist. Prof. Deepshikha Patel, Technocrat Institute of Technology, India  
Assist. Prof. Maram Balajee, GMRIT, India  
Assist. Prof. Monika Bhatnagar, TIT, India  
Prof. Gaurang Panchal, Charotar University of Science & Technology, India  
Prof. Anand K. Tripathi, Computer Society of India  
Prof. Jyoti Chaudhary, High Performance Computing Research Lab, India  
Assist. Prof. Supriya Raheja, ITM University, India

Dr. Pankaj Gupta, Microsoft Corporation, U.S.A.  
Assist. Prof. Panchamukesh Chandaka, Hyderabad Institute of Tech. & Management, India  
Prof. Mohan H.S, SJB Institute Of Technology, India  
Mr. Hossein Malekinezhad, Islamic Azad University, Iran  
Mr. Zatin Gupta, Universti Malaysia, Malaysia  
Assist. Prof. Amit Chauhan, Phonics Group of Institutions, India  
Assist. Prof. Ajal A. J., METS School Of Engineering, India  
Mrs. Omowunmi Omobola Adeyemo, University of Ibadan, Nigeria  
Dr. Bharat Bhushan Agarwal, I.F.T.M. University, India  
Md. Nazrul Islam, University of Western Ontario, Canada  
Tushar Kanti, L.N.C.T, Bhopal, India  
Er. Aumreesh Kumar Saxena, SIRTs College Bhopal, India  
Mr. Mohammad Monirul Islam, Daffodil International University, Bangladesh  
Dr. Kashif Nisar, University Utara Malaysia, Malaysia  
Dr. Wei Zheng, Rutgers Univ/ A10 Networks, USA  
Associate Prof. Rituraj Jain, Vyas Institute of Engg & Tech, Jodhpur – Rajasthan  
Assist. Prof. Apoorvi Sood, I.T.M. University, India  
Dr. Kayhan Zrar Ghafoor, University Technology Malaysia, Malaysia  
Mr. Swapnil Soner, Truba Institute College of Engineering & Technology, Indore, India  
Ms. Yogita Gigras, I.T.M. University, India  
Associate Prof. Neelima Sadineni, Pydha Engineering College, India Pydha Engineering College  
Assist. Prof. K. Deepika Rani, HITAM, Hyderabad  
Ms. Shikha Maheshwari, Jaipur Engineering College & Research Centre, India  
Prof. Dr V S Giridhar Akula, Avanthi's Scientific Tech. & Research Academy, Hyderabad  
Prof. Dr.S.Saravanan, Muthayammal Engineering College, India  
Mr. Mehdi Golsorkhtabar Amiri, Islamic Azad University, Iran  
Prof. Amit Sadanand Savyanavar, MITCOE, Pune, India  
Assist. Prof. P.Oliver Jayaprakash, Anna University,Chennai  
Assist. Prof. Ms. Sujata, ITM University, Gurgaon, India  
Dr. Asoke Nath, St. Xavier's College, India  
Mr. Masoud Rafighi, Islamic Azad University, Iran  
Assist. Prof. RamBabu Pemula, NIMRA College of Engineering & Technology, India  
Assist. Prof. Ms Rita Chhikara, ITM University, Gurgaon, India  
Mr. Sandeep Maan, Government Post Graduate College, India  
Prof. Dr. S. Muralidharan, Mepco Schlenk Engineering College, India  
Associate Prof. T.V.Sai Krishna, QIS College of Engineering and Technology, India  
Mr. R. Balu, Bharathiar University, Coimbatore, India  
Assist. Prof. Shekhar. R, Dr.SM College of Engineering, India  
Prof. P. Senthilkumar, Vivekananda Institue of Engineering and Techology for Woman, India  
Mr. M. Kamarajan, PSNA College of Engineering & Technology, India  
Dr. Angajala Srinivasa Rao, Jawaharlal Nehru Technical University, India  
Assist. Prof. C. Venkatesh, A.I.T.S, Rajampet, India  
Mr. Afshin Rezakhani Roozbahani, Ayatollah Boroujerdi University, Iran  
Mr. Laxmi chand, SCTL, Noida, India  
Dr. Dr. Abdul Hannan, Vivekanand College, Aurangabad  
Prof. Mahesh Panchal, KITRC, Gujarat  
Dr. A. Subramani, K.S.R. College of Engineering, Tiruchengode

Assist. Prof. Prakash M, Rajalakshmi Engineering College, Chennai, India  
Assist. Prof. Akhilesh K Sharma, Sir Padampat Singhania University, India  
Ms. Varsha Sahni, Guru Nanak Dev Engineering College, Ludhiana, India  
Associate Prof. Trilochan Rout, NM Institute of Engineering and Technology, India  
Mr. Srikanta Kumar Mohapatra, NMIET, Orissa, India  
Mr. Waqas Haider Bangyal, Iqra University Islamabad, Pakistan  
Dr. S. Vijayaragavan, Christ College of Engineering and Technology, Pondicherry, India  
Prof. Elboukhari Mohamed, University Mohammed First, Oujda, Morocco  
Dr. Muhammad Asif Khan, King Faisal University, Saudi Arabia  
Dr. Nagy Ramadan Darwish Omran, Cairo University, Egypt.  
Assistant Prof. Anand Nayyar, KCL Institute of Management and Technology, India  
Mr. G. Premsankar, Ericcson, India  
Assist. Prof. T. Hemalatha, VELS University, India  
Prof. Tejaswini Apte, University of Pune, India  
Dr. Edmund Ng Giap Weng, Universiti Malaysia Sarawak, Malaysia  
Mr. Mahdi Nouri, Iran University of Science and Technology, Iran  
Associate Prof. S. Asif Hussain, Annamacharya Institute of technology & Sciences, India  
Mrs. Kavita Pabreja, Maharaja Surajmal Institute (an affiliate of GGSIP University), India  
Mr. Vorugunti Chandra Sekhar, DA-IICT, India  
Mr. Muhammad Najmi Ahmad Zabidi, Universiti Teknologi Malaysia, Malaysia  
Dr. Aderemi A. Atayero, Covenant University, Nigeria  
Assist. Prof. Osama Sohaib, Balochistan University of Information Technology, Pakistan  
Assist. Prof. K. Suresh, Annamacharya Institute of Technology and Sciences, India  
Mr. Hassen Mohammed Abdullaah Alsafi, International Islamic University Malaysia (IIUM) Malaysia  
Mr. Robail Yasrab, Virtual University of Pakistan, Pakistan  
Mr. R. Balu, Bharathiar University, Coimbatore, India  
Prof. Anand Nayyar, KCL Institute of Management and Technology, Jalandhar  
Assoc. Prof. Vivek S Deshpande, MIT College of Engineering, India  
Prof. K. Saravanan, Anna university Coimbatore, India  
Dr. Ravendra Singh, MJP Rohilkhand University, Bareilly, India  
Mr. V. Mathivanan, IBRA College of Technology, Sultanate of OMAN  
Assoc. Prof. S. Asif Hussain, AITS, India  
Assist. Prof. C. Venkatesh, AITS, India  
Mr. Sami Ulhaq, SZABIST Islamabad, Pakistan  
Dr. B. Justus Rabi, Institute of Science & Technology, India  
Mr. Anuj Kumar Yadav, Dehradun Institute of technology, India  
Mr. Alejandro Mosquera, University of Alicante, Spain  
Assist. Prof. Arjun Singh, Sir Padampat Singhania University (SPSU), Udaipur, India  
Dr. Smriti Agrawal, JB Institute of Engineering and Technology, Hyderabad  
Assist. Prof. Swathi Sambangi, Visakha Institute of Engineering and Technology, India  
Ms. Prabhjot Kaur, Guru Gobind Singh Indraprastha University, India  
Mrs. Samaher AL-Hothali, Yanbu University College, Saudi Arabia  
Prof. Rajneeshkaur Bedi, MIT College of Engineering, Pune, India  
Mr. Hassen Mohammed Abdullaah Alsafi, International Islamic University Malaysia (IIUM)  
Dr. Wei Zhang, Amazon.com, Seattle, WA, USA  
Mr. B. Santhosh Kumar, C S I College of Engineering, Tamil Nadu  
Dr. K. Reji Kumar, , N S S College, Pandalam, India

Assoc. Prof. K. Seshadri Sastry, EIILM University, India  
Mr. Kai Pan, UNC Charlotte, USA  
Mr. Ruikar Sachin, SGGSIET, India  
Prof. (Dr.) Vinodani Katiyar, Sri Ramswaroop Memorial University, India  
Assoc. Prof., M. Giri, Sreenivasa Institute of Technology and Management Studies, India  
Assoc. Prof. Labib Francis Gergis, Misr Academy for Engineering and Technology (MET), Egypt  
Assist. Prof. Amanpreet Kaur, ITM University, India  
Assist. Prof. Anand Singh Rajawat, Shri Vaishnav Institute of Technology & Science, Indore  
Mrs. Hadeel Saleh Haj Aliwi, Universiti Sains Malaysia (USM), Malaysia  
Dr. Abhay Bansal, Amity University, India  
Dr. Mohammad A. Mezher, Fahad Bin Sultan University, KSA  
Assist. Prof. Nidhi Arora, M.C.A. Institute, India  
Prof. Dr. P. Suresh, Karpagam College of Engineering, Coimbatore, India  
Dr. Kannan Balasubramanian, Mepco Schlenk Engineering College, India  
Dr. S. Sankara Gomathi, Panimalar Engineering college, India  
Prof. Anil Kumar Suthar, Gujarat Technological University, L.C. Institute of Technology, India  
Assist. Prof. R. Hubert Rajan, NOORUL ISLAM UNIVERSITY, India  
Assist. Prof. Dr. Jyoti Mahajan, College of Engineering & Technology  
Assist. Prof. Homam Reda El-Taj, College of Network Engineering, Saudi Arabia & Malaysia  
Mr. Bijan Paul, Shahjalal University of Science & Technology, Bangladesh  
Assoc. Prof. Dr. Ch V Phani Krishna, KL University, India  
Dr. Vishal Bhatnagar, Ambedkar Institute of Advanced Communication Technologies & Research, India  
Dr. Lamri LAOUAMER, Al Qassim University, Dept. Info. Systems & European University of Brittany, Dept. Computer Science, UBO, Brest, France  
Prof. Ashish Babanrao Sasankar, G.H.Raisoni Institute Of Information Technology, India  
Prof. Pawan Kumar Goel, Shamli Institute of Engineering and Technology, India  
Mr. Ram Kumar Singh, S.V Subharti University, India  
Assistant Prof. Sunish Kumar O S, Amaljyothi College of Engineering, India  
Dr Sanjay Bhargava, Banasthali University, India  
Mr. Pankaj S. Kulkarni, AVEW's Shatabdi Institute of Technology, India  
Mr. Roohollah Etemadi, Islamic Azad University, Iran  
Mr. Oloruntoyin Sefiu Taiwo, Emmanuel Alayande College Of Education, Nigeria  
Mr. Sumit Goyal, National Dairy Research Institute, India  
Mr Jaswinder Singh Dilawari, Geeta Engineering College, India  
Prof. Raghuraj Singh, Harcourt Butler Technological Institute, Kanpur  
Dr. S.K. Mahendran, Anna University, Chennai, India  
Dr. Amit Wason, Hindustan Institute of Technology & Management, Punjab  
Dr. Ashu Gupta, Apeejay Institute of Management, India  
Assist. Prof. D. Asir Antony Gnana Singh, M.I.E.T Engineering College, India  
Mrs Mina Farmanbar, Eastern Mediterranean University, Famagusta, North Cyprus  
Mr. Maram Balajee, GMR Institute of Technology, India  
Mr. Moiz S. Ansari, Isra University, Hyderabad, Pakistan  
Mr. Adebayo, Olawale Surajudeen, Federal University of Technology Minna, Nigeria  
Mr. Jasvir Singh, University College Of Engg., India  
Mr. Vivek Tiwari, MANIT, Bhopal, India  
Assoc. Prof. R. Navaneethakrishnan, Bharathiyar College of Engineering and Technology, India  
Mr. Somdip Dey, St. Xavier's College, Kolkata, India

Mr. Souleymane Balla-Arabé, Xi'an University of Electronic Science and Technology, China  
Mr. Mahabub Alam, Rajshahi University of Engineering and Technology, Bangladesh  
Mr. Sathyaprakash P., S.K.P Engineering College, India  
Dr. N. Karthikeyan, SNS College of Engineering, Anna University, India  
Dr. Binod Kumar, JSPM's, Jayawant Technical Campus, Pune, India  
Assoc. Prof. Dinesh Goyal, Suresh Gyan Vihar University, India  
Mr. Md. Abdul Ahad, K L University, India  
Mr. Vikas Bajpai, The LNM IIT, India  
Dr. Manish Kumar Anand, Salesforce (R & D Analytics), San Francisco, USA  
Assist. Prof. Dheeraj Murari, Kumaon Engineering College, India  
Assoc. Prof. Dr. A. Muthukumaravel, VELS University, Chennai  
Mr. A. Siles Balasingh, St.Joseph University in Tanzania, Tanzania  
Mr. Ravindra Daga Badgujar, R C Patel Institute of Technology, India  
Dr. Preeti Khanna, SVKM's NMIMS, School of Business Management, India  
Mr. Kumar Dayanand, Cambridge Institute of Technology, India  
Dr. Syed Asif Ali, SMI University Karachi, Pakistan  
Prof. Pallvi Pandit, Himachal Pradesh University, India  
Mr. Ricardo Verschueren, University of Gloucestershire, UK  
Assist. Prof. Mamta Juneja, University Institute of Engineering and Technology, Panjab University, India  
Assoc. Prof. P. Surendra Varma, NRI Institute of Technology, JNTU Kakinada, India  
Assist. Prof. Gaurav Shrivastava, RGPV / SVITS Indore, India  
Dr. S. Sumathi, Anna University, India  
Assist. Prof. Ankita M. Kapadia, Charotar University of Science and Technology, India  
Mr. Deepak Kumar, Indian Institute of Technology (BHU), India  
Dr. Dr. Rajan Gupta, GGSIP University, New Delhi, India  
Assist. Prof M. Anand Kumar, Karpagam University, Coimbatore, India  
Mr. Mr Arshad Mansoor, Pakistan Aeronautical Complex  
Mr. Kapil Kumar Gupta, Ansal Institute of Technology and Management, India  
Dr. Neeraj Tomer, SINE International Institute of Technology, Jaipur, India  
Assist. Prof. Trunal J. Patel, C.G.Patel Institute of Technology, Uka Tarsadia University, Bardoli, Surat  
Mr. Sivakumar, Codework solutions, India  
Mr. Mohammad Sadegh Mirzaei, PGNR Company, Iran  
Dr. Gerard G. Dumancas, Oklahoma Medical Research Foundation, USA  
Mr. Varadala Sridhar, Varadhaman College Engineering College, Affiliated To JNTU, Hyderabad  
Assist. Prof. Manoj Dhawan, SVITS, Indore  
Assoc. Prof. Chitreshh Banerjee, Suresh Gyan Vihar University, Jaipur, India  
Dr. S. Santhi, SCSVMV University, India  
Mr. Davood Mohammadi Souran, Ministry of Energy of Iran, Iran  
Mr. Shamim Ahmed, Bangladesh University of Business and Technology, Bangladesh  
Mr. Sandeep Reddivari, Mississippi State University, USA  
Assoc. Prof. Ousmane Thiare, Gaston Berger University, Senegal  
Dr. Hazra Imran, Athabasca University, Canada  
Dr. Setu Kumar Chaturvedi, Technocrats Institute of Technology, Bhopal, India  
Mr. Mohd Dilshad Ansari, Jaypee University of Information Technology, India  
Ms. Jaspreet Kaur, Distance Education LPU, India  
Dr. D. Nagarajan, Salalah College of Technology, Sultanate of Oman  
Dr. K.V.N.R.Sai Krishna, S.V.R.M. College, India

- Mr. Himanshu Pareek, Center for Development of Advanced Computing (CDAC), India  
Mr. Khaldi Amine, Badji Mokhtar University, Algeria  
Mr. Mohammad Sadegh Mirzaei, Scientific Applied University, Iran  
Assist. Prof. Khyati Chaudhary, Ram-eesh Institute of Engg. & Technology, India  
Mr. Sanjay Agal, Pacific College of Engineering Udaipur, India  
Mr. Abdul Mateen Ansari, King Khalid University, Saudi Arabia  
Dr. H.S. Behera, Veer Surendra Sai University of Technology (VSSUT), India  
Dr. Shrikant Tiwari, Shri Shankaracharya Group of Institutions (SSGI), India  
Prof. Ganesh B. Regulwar, Shri Shankarprasad Agnihotri College of Engg, India  
Prof. Pinnamaneni Bhanu Prasad, Matrix vision GmbH, Germany  
Dr. Shrikant Tiwari, Shri Shankaracharya Technical Campus (SSTC), India  
Dr. Siddesh G.K., : Dayananada Sagar College of Engineering, Bangalore, India  
Dr. Nadir Bouchama, CERIST Research Center, Algeria  
Dr. R. Sathishkumar, Sri Venkateswara College of Engineering, India  
Assistant Prof (Dr.) Mohamed Moussaoui, Abdelmalek Essaadi University, Morocco  
Dr. S. Malathi, Panimalar Engineering College, Chennai, India  
Dr. V. Subedha, Panimalar Institute of Technology, Chennai, India  
Dr. Prashant Panse, Swami Vivekanand College of Engineering, Indore, India  
Dr. Hamza Aldabbas, Al-Balqa'a Applied University, Jordan  
Dr. G. Rasitha Banu, Vel's University, Chennai  
Dr. V. D. Ambeth Kumar, Panimalar Engineering College, Chennai  
Prof. Anuranjan Misra, Bhagwant Institute of Technology, Ghaziabad, India  
Ms. U. Sinthuja, PSG college of arts &science, India  
Dr. Ehsan Saradar Torshizi, Urmia University, Iran  
Dr. Shamneesh Sharma, APG Shimla University, Shimla (H.P.), India  
Assistant Prof. A. S. Syed Navaz, Muthayammal College of Arts & Science, India  
Assistant Prof. Ranjit Panigrahi, Sikkim Manipal Institute of Technology, Majitar, Sikkim  
Dr. Khaled Eskaf, Arab Academy for Science ,Technology & Maritime Transportation, Egypt  
Dr. Nishant Gupta, University of Jammu, India  
Assistant Prof. Nagarajan Sankaran, Annamalai University, Chidambaram, Tamilnadu, India  
Assistant Prof. Tribikram Pradhan, Manipal Institute of Technology, India  
Dr. Nasser Lotfi, Eastern Mediterranean University, Northern Cyprus  
Dr. R. Manavalan, K S Rangasamy college of Arts and Science, Tamilnadu, India  
Assistant Prof. P. Krishna Sankar, K S Rangasamy college of Arts and Science, Tamilnadu, India  
Dr. Rahul Malik, Cisco Systems, USA  
Dr. S. C. Lingareddy, ALPHA College of Engineering, India  
Assistant Prof. Mohammed Shuaib, Interl University, Lucknow, India  
Dr. Sachin Yele, Sanghvi Institute of Management & Science, India  
Dr. T. Thambidurai, Sun Univercell, Singapore  
Prof. Anandkumar Telang, BKIT, India  
Assistant Prof. R. Poorvadevi, SCSVMV University, India  
Dr Uttam Mande, Gitam University, India  
Dr. Poornima Girish Naik, Shahu Institute of Business Education and Research (SIBER), India  
Prof. Md. Abu Kausar, Jaipur National University, Jaipur, India  
Dr. Mohammed Zuber, AISECT University, India  
Prof. Kalum Priyanath Udagepola, King Abdulaziz University, Saudi Arabia  
Dr. K. R. Ananth, Velalar College of Engineering and Technology, India

Assistant Prof. Sanjay Sharma, Roorkee Engineering & Management Institute Shamli (U.P), India  
Assistant Prof. Panem Charan Arur, Priyadarshini Institute of Technology, India  
Dr. Ashwak Mahmood muhsen alabaichi, Karbala University / College of Science, Iraq  
Dr. Urmila Shrawankar, G H Raisoni College of Engineering, Nagpur (MS), India  
Dr. Krishan Kumar Paliwal, Panipat Institute of Engineering & Technology, India  
Dr. Mukesh Negi, Tech Mahindra, India  
Dr. Anuj Kumar Singh, Amity University Gurgaon, India  
Dr. Babar Shah, Gyeongsang National University, South Korea  
Assistant Prof. Jayprakash Upadhyay, SRI-TECH Jabalpur, India  
Assistant Prof. Varadala Sridhar, Vidya Jyothi Institute of Technology, India  
Assistant Prof. Parameshachari B D, KSIT, Bangalore, India  
Assistant Prof. Ankit Garg, Amity University, Haryana, India  
Assistant Prof. Rajashe Karappa, SDMCET, Karnataka, India  
Assistant Prof. Varun Jasuja, GNIT, India  
Assistant Prof. Sonal Honale, Abha Gaikwad Patil College of Engineering Nagpur, India  
Dr. Pooja Choudhary, CT Group of Institutions, NIT Jalandhar, India  
Dr. Faouzi Hidoussi, UHL Batna, Algeria  
Dr. Naseer Ali Husieen, Wasit University, Iraq  
Assistant Prof. Vinod Kumar Shukla, Amity University, Dubai  
Dr. Ahmed Farouk Metwaly, K L University  
Mr. Mohammed Noaman Murad, Cihan University, Iraq  
Dr. Suxing Liu, Arkansas State University, USA  
Dr. M. Gomathi, Velalar College of Engineering and Technology, India  
Assistant Prof. Sumardiono, College PGRI Blitar, Indonesia  
Dr. Latika Kharb, Jagan Institute of Management Studies (JIMS), Delhi, India  
Associate Prof. S. Raja, Pauls College of Engineering and Technology, Tamilnadu, India  
Assistant Prof. Seyed Reza Pakize, Shahid Sani High School, Iran  
Dr. Thiyyagu Nagaraj, University-INOU, India  
Assistant Prof. Noreen Sarai, Harare Institute of Technology, Zimbabwe  
Assistant Prof. Gajanand Sharma, Suresh Gyan Vihar University Jaipur, Rajasthan, India  
Assistant Prof. Mapari Vikas Prakash, Siddhant COE, Sudumbare, Pune, India  
Dr. Devesh Katiyar, Shri Ramswaroop Memorial University, India  
Dr. Shenshen Liang, University of California, Santa Cruz, US  
Assistant Prof. Mohammad Abu Omar, Limkokwing University of Creative Technology- Malaysia  
Mr. Snehasis Banerjee, Tata Consultancy Services, India  
Assistant Prof. Kibona Lusekelo, Ruaha Catholic University (RUCU), Tanzania  
Assistant Prof. Adib Kabir Chowdhury, University College Technology Sarawak, Malaysia  
Dr. Ying Yang, Computer Science Department, Yale University, USA  
Dr. Vinay Shukla, Institute Of Technology & Management, India  
Dr. Liviu Octavian Mafteiu-Scai, West University of Timisoara, Romania  
Assistant Prof. Rana Khudhair Abbas Ahmed, Al-Rafidain University College, Iraq  
Assistant Prof. Nitin A. Naik, S.R.T.M. University, India  
Dr. Timothy Powers, University of Hertfordshire, UK  
Dr. S. Prasath, Bharathiar University, Erode, India  
Dr. Ritu Shrivastava, SIRTS Bhopal, India  
Prof. Rohit Shrivastava, Mittal Institute of Technology, Bhopal, India  
Dr. Gianina Mihai, Dunarea de Jos" University of Galati, Romania

Assistant Prof. Ms. T. Kalai Selvi, Erode Sengunthar Engineering College, India  
Assistant Prof. Ms. C. Kavitha, Erode Sengunthar Engineering College, India  
Assistant Prof. K. Sinivasamoorthi, Erode Sengunthar Engineering College, India  
Assistant Prof. Mallikarjun C Sarsamba Bheemnna Khandre Institute Technology, Bhalki, India  
Assistant Prof. Vishwanath Chikaraddi, Veermata Jijabai technological Institute (Central Technological Institute), India  
Assistant Prof. Dr. Ikvinderpal Singh, Trai Shatabdi GGS Khalsa College, India  
Assistant Prof. Mohammed Noaman Murad, Cihan University, Iraq  
Professor Yousef Farhaoui, Moulay Ismail University, Errachidia, Morocco  
Dr. Parul Verma, Amity University, India  
Professor Yousef Farhaoui, Moulay Ismail University, Errachidia, Morocco  
Assistant Prof. Madhavi Dhingra, Amity University, Madhya Pradesh, India  
Assistant Prof.. G. Selvavinayagam, SNS College of Technology, Coimbatore, India  
Assistant Prof. Madhavi Dhingra, Amity University, MP, India  
Professor Kartheesan Log, Anna University, Chennai  
Professor Vasudeva Acharya, Shri Madhwa vadira Institute of Technology, India  
Dr. Asif Iqbal Hajamydeen, Management & Science University, Malaysia  
Assistant Prof., Mahendra Singh Meena, Amity University Haryana  
Assistant Professor Manjeet Kaur, Amity University Haryana  
Dr. Mohamed Abd El-Basset Matwalli, Zagazig University, Egypt  
Dr. Ramani Kannan, Universiti Teknologi PETRONAS, Malaysia  
Assistant Prof. S. Jagadeesan Subramaniam, Anna University, India  
Assistant Prof. Dharmendra Choudhary, Tripura University, India  
Assistant Prof. Deepika Vodnala, SR Engineering College, India  
Dr. Kai Cong, Intel Corporation & Computer Science Department, Portland State University, USA  
Dr. Kailas R Patil, Vishwakarma Institute of Information Technology (VIIT), India  
Dr. Omar A. Alzubi, Faculty of IT / Al-Balqa Applied University, Jordan  
Assistant Prof. Kareemullah Shaik, Nimra Institute of Science and Technology, India  
Assistant Prof. Chirag Modi, NIT Goa  
Dr. R. Ramkumar, Nandha Arts And Science College, India  
Dr. Priyadarshini Vydhalingam, Harathiar University, India  
Dr. P. S. Jagadeesh Kumar, DBIT, Bangalore, Karnataka  
Dr. Vikas Thada, AMITY University, Pachgaon  
Dr. T. A. Ashok Kumar, Institute of Management, Christ University, Bangalore  
Dr. Shaheera Rashwan, Informatics Research Institute  
Dr. S. Preetha Gunasekar, Bharathiyar University, India  
Asst Professor Sameer Dev Sharma, Uttarakhand University, Dehradun  
Dr. Zhihan Lv, Chinese Academy of Science, China  
Dr. Ikvinderpal Singh, Trai Shatabdi GGS Khalsa College, Amritsar  
Dr. Umar Ruhi, University of Ottawa, Canada  
Dr. Jasmin Cosic, University of BiHac, Bosnia and Herzegovina  
Dr. Homam Reda El-Taj, University of Tabuk, Kingdom of Saudi Arabia  
Dr. Mostafa Ghobaei Arani, Islamic Azad University, Iran  
Dr. Ayyasamy Ayyanar, Annamalai University, India  
Dr. Selvakumar Manickam, Universiti Sains Malaysia, Malaysia  
Dr. Murali Krishna Namana, GITAM University, India  
Dr. Smriti Agrawal, Chaitanya Bharathi Institute of Technology, Hyderabad, India  
Professor Vimalathithan Rathinasabapathy, Karpagam College Of Engineering, India

Dr. Sushil Chandra Dimri, Graphic Era University, India  
Dr. Dinh-Sinh Mai, Le Quy Don Technical University, Vietnam  
Dr. S. Rama Sree, Aditya Engg. College, India  
Dr. Ehab T. Alnfrawy, Sadat Academy, Egypt  
Dr. Patrick D. Cerna, Haramaya University, Ethiopia  
Dr. Vishal Jain, Bharati Vidyapeeth's Institute of Computer Applications and Management (BVICAM), India  
Associate Prof. Dr. Jiliang Zhang, North Eastern University, China  
Dr. Sharefa Murad, Middle East University, Jordan  
Dr. Ajeet Singh Poonia, Govt. College of Engineering & technology, Rajasthan, India  
Dr. Vahid Esmaeelzadeh, University of Science and Technology, Iran  
Dr. Jacek M. Czerniak, Casimir the Great University in Bydgoszcz, Institute of Technology, Poland  
Associate Prof. Anisur Rehman Nasir, Jamia Millia Islamia University  
Assistant Prof. Imran Ahmad, COMSATS Institute of Information Technology, Pakistan  
Professor Ghulam Qasim, Preston University, Islamabad, Pakistan  
Dr. Parameshachari B D, GSSS Institute of Engineering and Technology for Women  
Dr. Wencan Luo, University of Pittsburgh, US  
Dr. Musa PEKER, Faculty of Technology, Mugla Sitki Kocman University, Turkey  
Dr. Gunasekaran Shanmugam, Anna University, India  
Dr. Binh P. Nguyen, National University of Singapore, Singapore  
Dr. Rajkumar Jain, Indian Institute of Technology Indore, India  
Dr. Imtiaz Ali Halepoto, QUEST Nawabshah, Pakistan  
Dr. Shaligram Prajapat, Devi Ahilya University Indore India  
Dr. Sunita Singhal, Birla Institute of Technology and Science, Pilani, India  
Dr. Ijaz Ali Shoukat, King Saud University, Saudi Arabia  
Dr. Anuj Gupta, IKG Punjab Technical University, India  
Dr. Sonali Saini, IES-IPS Academy, India  
Dr. Krishan Kumar, MotiLal Nehru National Institute of Technology, Allahabad, India  
Dr. Z. Faizal Khan, College of Engineering, Shaqra University, Kingdom of Saudi Arabia  
Prof. M. Padmavathamma, S.V. University Tirupati, India  
Prof. A. Velayudham, Cape Institute of Technology, India  
Prof. Seifeide Kadry, American University of the Middle East  
Dr. J. Durga Prasad Rao, Pt. Ravishankar Shukla University, Raipur  
Assistant Prof. Najam Hasan, Dhofar University  
Dr. G. Suseendran, Vels University, Pallavaram, Chennai  
Prof. Ankit Faldu, Gujarat Technological University- Atmiya Institute of Technology and Science  
Dr. Ali Habiboghi, Islamic Azad University  
Dr. Deepak Dembla, JECRC University, Jaipur, India  
Dr. Pankaj Rajan, Walmart Labs, USA  
Assistant Prof. Radoslava Kraleva, South-West University "Neofit Rilski", Bulgaria  
Assistant Prof. Medhavi Shriwas, Shri Vaishnav Institute of Technology, India  
Associate Prof. Sedat Akleylek, Ondokuz Mayis University, Turkey  
Dr. U.V. Arivazhagu, Kingston Engineering College Affiliated To Anna University, India  
Dr. Touseef Ali, University of Engineering and Technology, Taxila, Pakistan  
Assistant Prof. Naren Jeeva, SASTRA University, India  
Dr. Riccardo Colella, University of Salento, Italy  
Dr. Enache Maria Cristina, University of Galati, Romania  
Dr. Senthil P, Kurinji College of Arts & Science, India

- Dr. Hasan Ashrafi-rizi, Isfahan University of Medical Sciences, Isfahan, Iran  
Dr. Mazhar Malik, Institute of Southern Punjab, Pakistan  
Dr. Yajie Miao, Carnegie Mellon University, USA  
Dr. Kamran Shaukat, University of the Punjab, Pakistan  
Dr. Sasikaladevi N., SASTRA University, India  
Dr. Ali Asghar Rahmani Hosseiniabadi, Islamic Azad University Ayatollah Amoli Branch, Amol, Iran  
Dr. Velin Kralev, South-West University "Neofit Rilski", Blagoevgrad, Bulgaria  
Dr. Marius Iulian Mihailescu, LUMINA - The University of South-East Europe  
Dr. Sriramula Nagaprasad, S.R.R.Govt.Arts & Science College, Karimnagar, India  
Prof (Dr.) Namrata Dhanda, Dr. APJ Abdul Kalam Technical University, Lucknow, India  
Dr. Javed Ahmed Maher, Shah Abdul Latif University, Khairpur Mir's, Pakistan  
Dr. B. Narendra Kumar Rao, Sree Vidyanikethan Engineering College, India  
Dr. Shahzad Anwar, University of Engineering & Technology Peshawar, Pakistan  
Dr. Basit Shahzad, King Saud University, Riyadh - Saudi Arabia  
Dr. Nilamadhab Mishra, Chang Gung University  
Dr. Sachin Kumar, Indian Institute of Technology Roorkee  
Dr. Santosh Nanda, Biju-Patnaik University of Technology  
Dr. Sherzod Turaev, International Islamic University Malaysia  
Dr. Yilun Shang, Tongji University, Department of Mathematics, Shanghai, China  
Dr. Nuzhat Shaikh, Modern Education society's College of Engineering, Pune, India  
Dr. Parul Verma, Amity University, Lucknow campus, India  
Dr. Rachid Alaoui, Agadir Ibn Zohr University, Agadir, Morocco  
Dr. Dharmendra Patel, Charotar University of Science and Technology, India  
Dr. Dong Zhang, University of Central Florida, USA  
Dr. Kennedy Chinedu Okafor, Federal University of Technology Owerri, Nigeria  
Prof. C Ram Kumar, Dr NGP Institute of Technology, India  
Dr. Sandeep Gupta, GGS IP University, New Delhi, India  
Dr. Shahanawaj Ahamad, University of Ha'il, Ha'il City, Ministry of Higher Education, Kingdom of Saudi Arabia  
Dr. Najeed Ahmed Khan, NED University of Engineering & Technology, India  
Dr. Sajid Ullah Khan, Universiti Malaysia Sarawak, Malaysia  
Dr. Muhammad Asif, National Textile University Faisalabad, Pakistan  
Dr. Yu BI, University of Central Florida, Orlando, FL, USA  
Dr. Brijendra Kumar Joshi, Research Center, Military College of Telecommunication Engineering, India  
Prof. Dr. Nak Eun Cho, Pukyong National University, Korea  
Prof. Wasim Ul-Haq, Faculty of Science, Majmaah University, Saudi Arabia  
Dr. Mohsan Raza, G.C University Faisalabad, Pakistan  
Dr. Syed Zakar Hussain Bukhari, National Science and Technalogy Azad Jamu Kashmir, Pakistan  
Dr. Ruksar Fatima, KBN College of Engineering, Gulbarga, Karnataka, India  
Associate Professor S. Karpagavalli, Department of Computer Science, PSGR Krishnammal College for Women  
Coimbatore, Tamilnadu, India  
Dr. Bushra Mohamed Elamin Elnaim, Prince Sattam bin Abdulaziz University, Saudi Arabia  
Dr. Shamik Tiwari, Department of CSE, CET, Mody University, Lakshmangarh  
Dr. Rohit Raja, Faculty of Engineering and Technology, Shri Shankaracharya Group of Institutions, India  
Prof. Dr. Aqeel-ur-Rehman, Department of Computing, HIET, FEST, Hamdard University, Pakistan  
Dr. Nageswara Rao Moparthi, Velagapudi Ramakrishna Siddhartha Engineering College, India  
Dr. Mohd Muqeem, Department of Computer Application, Integral University, Lucknow, India  
Dr. Zeeshan Bhatti, Institute of Information and Communication Technology, University of Sindh, Jamshoro, Pakistan

- Dr. Emrah Irmak, Biomedical Engineering Department, Karabuk University, Turkey  
Dr. Fouad Abdulameer salman, School of Informatics and Applied Mathematics, Universiti Malaysia Terengganu  
Dr. N. Prasath, Department of Computer Science and Engineering, KPR Institute of Engineering and Technology, Arasur, Coimbatore  
Dr. Hasan Ashrafi-rizi, Health Information Technology Research Center, Isfahan University of Medical Sciences, Hezar Jerib Avenue, Isfahan, Iran  
Dr. N. Sasikaladevi, School of Computing, SASTRA University, Thirumalisamudram, Tamilnadu, India.  
Dr. Anchit Bijalwan, Arba Minch University, Ethiopia  
Dr. K. Sathishkumar, BlueCrest University College, Accra North, Ghana, West Africa  
Dr. Dr. Parameshachari B D, GSSS Institute of Engineering and Technology for Women, Affiliated to Visvesvaraya Technological University, Belagavi  
Dr. C. Shoba Bindu, Dept. of CSE, JNTUA College of Engineering, India  
Dr. M. Inbavalli, ER. Perumal Manimekalai College of Engineering, Hosur, Tamilnadu, India  
Dr. Vidya Sagar Ponnam, Dept. of IT, Velagapudi Ramakrishna Siddhartha Engineering College, India  
Dr. Kelvin LO M. F., The Hong Kong Polytechnic University, Hong Kong  
Prof. Karimella Vikram, G.H. Raisoni College of Engineering & Management, Pune, India  
Dr. Shajilin Loret J.B., VV College of Engineering, India  
Dr. P. Sujatha, Department of Computer Science at Vels University, Chennai  
Dr. Vaibhav Sundriyal, Old Dominion University Research Foundation, USA  
Dr. Md Masud Rana, Khulna University of Engineering and Technology, Bangladesh  
Dr. Gurcharan Singh, Khasa College Amritsar, Guru Nanak Dev University, Amritsar, India  
Dr. Richard Otieno Omollo, Department of Computer Science and Software Engineering, Jaramogi Oginga Odinga University of Science and Technology, Kenya  
Prof. (Dr) Amit Verma, Computer Science & Engineering, Chandigarh Engineering College, Landran, Mohali, India  
Dr. Vidya Sagar Ponnam, Velagapudi Ramakrishna Siddhartha Engineering College, India  
Dr. Bohui Wang, School of Aerospace Science and Technology, Xidian University, P.R. China  
Dr. M. Anjan Kumar, Department of Computer Science, Satavahana University, Karimnagar  
Dr. Hanumanthappa J., DoS in CS, Uni of Mysuru, Karnataka, India  
Dr. Pouya Derakhshan-Barjoei, Dept. of Telecommunication and Engineering, Islamic Azad University, Iran  
Dr. Tanweer Alam, Islamic University of Madinah, Dept. of Computer Science, College of Computer and Information System, Al Madinah, Saudi Arabia  
Dr. Kumar Keshamoni, Dept. of ECE, Vaagdevi Engineering College, Warangal, Telangana, India  
Dr. G. Rajkumar, N.M.S.S.Vellaichamy Nadar College, Madurai, Tamilnadu, India  
Dr. P. Mayil Vel Kumar, Karpagam Institute of Technology, Coimbatore, India  
Dr. M. Yaswanth Bhanu Murthy, Vasireddy Venkatadri Institute of Technology, Guntur, A.P., India  
Asst. Prof. Dr. Mehmet Barış TABAKCIOĞLU, Bursa Technical University, Turkey  
Dr. Mohd. Muntjir, College of Computers and Information Technology, Taif University, Kingdom of Saudi Arabia  
Dr. Sanjay Agal, Aravalı Institute of Technical Studies, Udaipur, India  
Dr. Shanshan Tuo, xAd Inc., US  
Dr. Subhadra Shaw, AKS University, Satna, India  
Dr. Piyush Anand, Noida International University, Greater Noida, India  
Dr. Brijendra Kumar Joshi, Research Center Military College of Telecommunication Engineering, India  
Dr. V. Sreerama Murthy, GMRIT, Rajam, AP, India  
Dr. S. Nagarajan, Annamalai University, India  
Prof. Pramod Bhausaheb Deshmukh, D. Y. Patil College of Engineering, Akurdi, Pune, India

# **CALL FOR PAPERS**

## **International Journal of Computer Science and Information Security**

**IJCSIS 2018-2019**

**ISSN: 1947-5500**

**<http://sites.google.com/site/ijcsis/>**

International Journal Computer Science and Information Security, IJCSIS, is the premier scholarly venue in the areas of computer science and security issues. IJCSIS 2011 will provide a high profile, leading edge platform for researchers and engineers alike to publish state-of-the-art research in the respective fields of information technology and communication security. The journal will feature a diverse mixture of publication articles including core and applied computer science related topics.

Authors are solicited to contribute to the special issue by submitting articles that illustrate research results, projects, surveying works and industrial experiences that describe significant advances in the following areas, but are not limited to. Submissions may span a broad range of topics, e.g.:

### ***Track A: Security***

Access control, Anonymity, Audit and audit reduction & Authentication and authorization, Applied cryptography, Cryptanalysis, Digital Signatures, Biometric security, Boundary control devices, Certification and accreditation, Cross-layer design for security, Security & Network Management, Data and system integrity, Database security, Defensive information warfare, Denial of service protection, Intrusion Detection, Anti-malware, Distributed systems security, Electronic commerce, E-mail security, Spam, Phishing, E-mail fraud, Virus, worms, Trojan Protection, Grid security, Information hiding and watermarking & Information survivability, Insider threat protection, Integrity  
Intellectual property protection, Internet/Intranet Security, Key management and key recovery, Language-based security, Mobile and wireless security, Mobile, Ad Hoc and Sensor Network Security, Monitoring and surveillance, Multimedia security ,Operating system security, Peer-to-peer security, Performance Evaluations of Protocols & Security Application, Privacy and data protection, Product evaluation criteria and compliance, Risk evaluation and security certification, Risk/vulnerability assessment, Security & Network Management, Security Models & protocols, Security threats & countermeasures (DDoS, MiM, Session Hijacking, Replay attack etc.,) Trusted computing, Ubiquitous Computing Security, Virtualization security, VoIP security, Web 2.0 security, Submission Procedures, Active Defense Systems, Adaptive Defense Systems, Benchmark, Analysis and Evaluation of Security Systems, Distributed Access Control and Trust Management, Distributed Attack Systems and Mechanisms, Distributed Intrusion Detection/Prevention Systems, Denial-of-Service Attacks and Countermeasures, High Performance Security Systems, Identity Management and Authentication, Implementation, Deployment and Management of Security Systems, Intelligent Defense Systems, Internet and Network Forensics, Large-scale Attacks and Defense, RFID Security and Privacy, Security Architectures in Distributed Network Systems, Security for Critical Infrastructures, Security for P2P systems and Grid Systems, Security in E-Commerce, Security and Privacy in Wireless Networks, Secure Mobile Agents and Mobile Code, Security Protocols, Security Simulation and Tools, Security Theory and Tools, Standards and Assurance Methods, Trusted Computing, Viruses, Worms, and Other Malicious Code, World Wide Web Security, Novel and emerging secure architecture, Study of attack strategies, attack modeling, Case studies and analysis of actual attacks, Continuity of Operations during an attack, Key management, Trust management, Intrusion detection techniques, Intrusion response, alarm management, and correlation analysis, Study of tradeoffs between security and system performance, Intrusion tolerance systems, Secure protocols, Security in wireless networks (e.g. mesh networks, sensor networks, etc.), Cryptography and Secure Communications, Computer Forensics, Recovery and Healing, Security Visualization, Formal Methods in Security, Principles for Designing a Secure Computing System, Autonomic Security, Internet Security, Security in Health Care Systems, Security Solutions Using Reconfigurable Computing, Adaptive and Intelligent Defense Systems, Authentication and Access control, Denial of service attacks and countermeasures, Identity, Route and

Location Anonymity schemes, Intrusion detection and prevention techniques, Cryptography, encryption algorithms and Key management schemes, Secure routing schemes, Secure neighbor discovery and localization, Trust establishment and maintenance, Confidentiality and data integrity, Security architectures, deployments and solutions, Emerging threats to cloud-based services, Security model for new services, Cloud-aware web service security, Information hiding in Cloud Computing, Securing distributed data storage in cloud, Security, privacy and trust in mobile computing systems and applications, **Middleware security & Security features:** middleware software is an asset on

its own and has to be protected, interaction between security-specific and other middleware features, e.g., context-awareness, **Middleware-level security monitoring and measurement:** metrics and mechanisms for quantification and evaluation of security enforced by the middleware, **Security co-design:** trade-off and co-design between application-based and middleware-based security, **Policy-based management:** innovative support for policy-based definition and enforcement of security concerns, **Identification and authentication mechanisms:** Means to capture application specific constraints in defining and enforcing access control rules, **Middleware-oriented security patterns:** identification of patterns for sound, reusable security, **Security in aspect-based middleware:** mechanisms for isolating and enforcing security aspects, **Security in agent-based platforms:** protection for mobile code and platforms, Smart Devices: Biometrics, National ID cards, Embedded Systems Security and TPMs, RFID Systems Security, Smart Card Security, Pervasive Systems: Digital Rights Management (DRM) in pervasive environments, Intrusion Detection and Information Filtering, Localization Systems Security (Tracking of People and Goods), Mobile Commerce Security, Privacy Enhancing Technologies, Security Protocols (for Identification and Authentication, Confidentiality and Privacy, and Integrity), Ubiquitous Networks: Ad Hoc Networks Security, Delay-Tolerant Network Security, Domestic Network Security, Peer-to-Peer Networks Security, Security Issues in Mobile and Ubiquitous Networks, Security of GSM/GPRS/UMTS Systems, Sensor Networks Security, Vehicular Network Security, Wireless Communication Security: Bluetooth, NFC, WiFi, WiMAX, WiMedia, others

This Track will emphasize the design, implementation, management and applications of computer communications, networks and services. Topics of mostly theoretical nature are also welcome, provided there is clear practical potential in applying the results of such work.

#### **Track B: Computer Science**

Broadband wireless technologies: LTE, WiMAX, WiRAN, HSDPA, HSUPA, Resource allocation and interference management, Quality of service and scheduling methods, Capacity planning and dimensioning, Cross-layer design and Physical layer based issue, Interworking architecture and interoperability, Relay assisted and cooperative communications, Location and provisioning and mobility management, Call admission and flow/congestion control, Performance optimization, Channel capacity modeling and analysis, Middleware Issues: Event-based, publish/subscribe, and message-oriented middleware, Reconfigurable, adaptable, and reflective middleware approaches, Middleware solutions for reliability, fault tolerance, and quality-of-service, Scalability of middleware, Context-aware middleware, Autonomic and self-managing middleware, Evaluation techniques for middleware solutions, Formal methods and tools for designing, verifying, and evaluating, middleware, Software engineering techniques for middleware, Service oriented middleware, Agent-based middleware, Security middleware, Network Applications: Network-based automation, Cloud applications, Ubiquitous and pervasive applications, Collaborative applications, RFID and sensor network applications, Mobile applications, Smart home applications, Infrastructure monitoring and control applications, Remote health monitoring, GPS and location-based applications, Networked vehicles applications, Alert applications, Embedded Computer System, Advanced Control Systems, and Intelligent Control : Advanced control and measurement, computer and microprocessor-based control, signal processing, estimation and identification techniques, application specific IC's, nonlinear and adaptive control, optimal and robot control, intelligent control, evolutionary computing, and intelligent systems, instrumentation subject to critical conditions, automotive, marine and aero-space control and all other control applications, Intelligent Control System, Wiring/Wireless Sensor, Signal Control System. Sensors, Actuators and Systems Integration : Intelligent sensors and actuators, multisensor fusion, sensor array and multi-channel processing, micro/nano technology, microsensors and microactuators, instrumentation electronics, MEMS and system integration, wireless sensor, Network Sensor, Hybrid

Sensor, Distributed Sensor Networks. Signal and Image Processing : Digital signal processing theory, methods, DSP implementation, speech processing, image and multidimensional signal processing, Image analysis and processing, Image and Multimedia applications, Real-time multimedia signal processing, Computer vision, Emerging signal processing areas, Remote Sensing, Signal processing in education. Industrial Informatics: Industrial applications of neural networks, fuzzy algorithms, Neuro-Fuzzy application, bioInformatics, real-time computer control, real-time information systems, human-machine interfaces, CAD/CAM/CAT/CIM, virtual reality, industrial communications, flexible manufacturing systems, industrial automated process, Data Storage Management, Harddisk control, Supply Chain Management, Logistics applications, Power plant automation, Drives automation. Information Technology, Management of Information System : Management information systems, Information Management, Nursing information management, Information System, Information Technology and their application, Data retrieval, Data Base Management, Decision analysis methods, Information processing, Operations research, E-Business, E-Commerce, E-Government, Computer Business, Security and risk management, Medical imaging, Biotechnology, Bio-Medicine, Computer-based information systems in health care, Changing Access to Patient Information, Healthcare Management Information Technology. Communication/Computer Network, Transportation Application : On-board diagnostics, Active safety systems, Communication systems, Wireless technology, Communication application, Navigation and Guidance, Vision-based applications, Speech interface, Sensor fusion, Networking theory and technologies, Transportation information, Autonomous vehicle, Vehicle application of affective computing, Advance Computing technology and their application : Broadband and intelligent networks, Data Mining, Data fusion, Computational intelligence, Information and data security, Information indexing and retrieval, Information processing, Information systems and applications, Internet applications and performances, Knowledge based systems, Knowledge management, Software Engineering, Decision making, Mobile networks and services, Network management and services, Neural Network, Fuzzy logics, Neuro-Fuzzy, Expert approaches, Innovation Technology and Management : Innovation and product development, Emerging advances in business and its applications, Creativity in Internet management and retailing, B2B and B2C management, Electronic transceiver device for Retail Marketing Industries, Facilities planning and management, Innovative pervasive computing applications, Programming paradigms for pervasive systems, Software evolution and maintenance in pervasive systems, Middleware services and agent technologies, Adaptive, autonomic and context-aware computing, Mobile/Wireless computing systems and services in pervasive computing, Energy-efficient and green pervasive computing, Communication architectures for pervasive computing, Ad hoc networks for pervasive communications, Pervasive opportunistic communications and applications, Enabling technologies for pervasive systems (e.g., wireless BAN, PAN), Positioning and tracking technologies, Sensors and RFID in pervasive systems, Multimodal sensing and context for pervasive applications, Pervasive sensing, perception and semantic interpretation, Smart devices and intelligent environments, Trust, security and privacy issues in pervasive systems, User interfaces and interaction models, Virtual immersive communications, Wearable computers, Standards and interfaces for pervasive computing environments, Social and economic models for pervasive systems, Active and Programmable Networks, Ad Hoc & Sensor Network, Congestion and/or Flow Control, Content Distribution, Grid Networking, High-speed Network Architectures, Internet Services and Applications, Optical Networks, Mobile and Wireless Networks, Network Modeling and Simulation, Multicast, Multimedia Communications, Network Control and Management, Network Protocols, Network Performance, Network Measurement, Peer to Peer and Overlay Networks, Quality of Service and Quality of Experience, Ubiquitous Networks, Crosscutting Themes – Internet Technologies, Infrastructure, Services and Applications; Open Source Tools, Open Models and Architectures; Security, Privacy and Trust; Navigation Systems, Location Based Services; Social Networks and Online Communities; ICT Convergence, Digital Economy and Digital Divide, Neural Networks, Pattern Recognition, Computer Vision, Advanced Computing Architectures and New Programming Models, Visualization and Virtual Reality as Applied to Computational Science, Computer Architecture and Embedded Systems, Technology in Education, Theoretical Computer Science, Computing Ethics, Computing Practices & Applications

Authors are invited to submit papers through e-mail [ijcsiseditor@gmail.com](mailto:ijcsiseditor@gmail.com). Submissions must be original and should not have been published previously or be under consideration for publication while being evaluated by IJCSIS. Before submission authors should carefully read over the journal's Author Guidelines, which are located at <http://sites.google.com/site/ijcsis/authors-notes>.

**© IJCSIS PUBLICATION 2018  
ISSN 1947 5500  
<http://sites.google.com/site/ijcsis/>**