Online ISSN : 0975-4172
Print ISSN : 0975-4350
DOI : 10.17406/GJCST

Global Journal

of Computer Science and Technology: C

Software & Data Engineering

Neuro Fuzzy Algorithm

Multi-Instance Multi-Cluste

Highlights

Software Effort Estimation

Cluster Analysis of Medical

Discovering Thoughts, Inventing Future

VOLUME 16 ISSUE 1 VERSION 1.0

Global Journal of Computer Science and Technology: C
Software & Data Engineering

Global Journal of Computer Science and Technology: C
Software & Data Engineering

Volume 16 Issue 1 (Ver. 1.0)

Open Association of Research Society

Incorporation No.: 0423089
License No.: 42125/022010/1186
Registration No.: 430374
Import-Export Code: 1109007027
Employer Identification Number (EIN):
USA Tax ID: 98-0673427

Global Journals Inc.

*(A Delaware USA Incorporation with "Good Standing"; **Reg. Number: 0423089**)*

Sponsors: *Open Association of Research Society*
Open Scientific Standards

Publisher's Headquarters office

Global Journals® Headquarters
945th Concord Streets,
Framingham Massachusetts Pin: 01701,
United States of America
USA Toll Free: +001-888-839-7392
USA Toll Free Fax: +001-888-839-7392

Offset Typesetting

Global Journals Incorporated
2nd, Lansdowne, Lansdowne Rd., Croydon-Surrey,
Pin: CR9 2ER, United Kingdom

Packaging & Continental Dispatching

Global Journals
E-3130 Sudama Nagar, Near Gopur Square,
Indore, M.P., Pin: 452009, India

Find a correspondence nodal officer near you

To find nodal officer of your country, please email us at *local@globaljournals.org*

eContacts

Press Inquiries: *press@globaljournals.org*
Investor Inquiries: *investors@globaljournals.org*
Technical Support: *technology@globaljournals.org*
Media & Releases: *media@globaljournals.org*

Pricing (Including by Air Parcel Charges):

For Authors:
22 USD (B/W) & 50 USD (Color)
Yearly Subscription (Personal & Institutional):
200 USD (B/W) & 250 USD (Color)

Integrated Editorial Board
(Computer Science, Engineering, Medical, Management, Natural Science, Social Science)

Dr. Bart Lambrecht
Director of Research in Accounting and Finance Professor of Finance
Lancaster University Management School
BA (Antwerp); MPhil, MA, PhD (Cambridge)

Dr. Carlos García Pont
Associate Professor of Marketing
IESE Business School, University of Navarra
Doctor of Philosophy (Management), Massachusetts Institute of Technology (MIT)
Master in Business Administration, IESE, University of Navarra
Degree in Industrial Engineering, Universitat Politècnica de Catalunya

Dr. Fotini Labropulu
Mathematics - Luther College
University of Regina Ph.D., M.Sc. in Mathematics
B.A. (Honors) in Mathematics
University of Windso

Dr. Lynn Lim
Reader in Business and Marketing
Roehampton University, London
BCom, PGDip, MBA (Distinction), PhD, FHEA

Dr. Mihaly Mezei
ASSOCIATE PROFESSOR
Department of Structural and Chemical Biology, Mount Sinai School of Medical Center
Ph.D., Etvs Lornd University
Postdoctoral Training,
New York University

Dr. Söhnke M. Bartram
Department of Accounting and Finance Lancaster University Management School Ph.D. (WHU Koblenz)
MBA/BBA (University of Saarbrücken)

Dr. Miguel Angel Ariño
Professor of Decision Sciences
IESE Business School
Barcelona, Spain (Universidad de Navarra)
CEIBS (China Europe International Business School).
Beijing, Shanghai and Shenzhen
Ph.D. in Mathematics
University of Barcelona
BA in Mathematics (Licenciatura)
University of Barcelona

Philip G. Moscoso
Technology and Operations Management
IESE Business School, University of Navarra
Ph.D in Industrial Engineering and Management, ETH Zurich
M.Sc. in Chemical Engineering, ETH Zurich

Dr. Sanjay Dixit, M.D.
Director, EP Laboratories, Philadelphia VA Medical Center
Cardiovascular Medicine - Cardiac Arrhythmia
Univ of Penn School of Medicine

Dr. Han-Xiang Deng
MD., Ph.D
Associate Professor and Research Department Division of Neuromuscular Medicine
Davee Department of Neurology and Clinical Neuroscience Northwestern University Feinberg School of Medicine

Dr. Pina C. Sanelli
Associate Professor of Public Health
Weill Cornell Medical College
Associate Attending Radiologist
NewYork-Presbyterian Hospital
MRI, MRA, CT, and CTA
Neuroradiology and Diagnostic Radiology
M.D., State University of New York at Buffalo, School of Medicine and Biomedical Sciences

Dr. Roberto Sanchez
Associate Professor
Department of Structural and Chemical Biology
Mount Sinai School of Medicine
Ph.D., The Rockefeller University

Dr. Wen-Yih Sun
Professor of Earth and Atmospheric Sciences Purdue University Director National Center for Typhoon and Flooding Research, Taiwan
University Chair Professor
Department of Atmospheric Sciences, National Central University, Chung-Li, Taiwan University Chair Professor
Institute of Environmental Engineering, National Chiao Tung University, Hsin-chu, Taiwan.Ph.D., MS The University of Chicago, Geophysical Sciences
BS National Taiwan University, Atmospheric Sciences
Associate Professor of Radiology

Dr. Michael R. Rudnick
M.D., FACP
Associate Professor of Medicine
Chief, Renal Electrolyte and Hypertension Division (PMC)
Penn Medicine, University of Pennsylvania
Presbyterian Medical Center, Philadelphia
Nephrology and Internal Medicine
Certified by the American Board of Internal Medicine

Dr. Bassey Benjamin Esu
B.Sc. Marketing; MBA Marketing; Ph.D Marketing
Lecturer, Department of Marketing, University of Calabar
Tourism Consultant, Cross River State Tourism Development Department
Co-ordinator , Sustainable Tourism Initiative, Calabar, Nigeria

Dr. Aziz M. Barbar, Ph.D.
IEEE Senior Member
Chairperson, Department of Computer Science
AUST - American University of Science & Technology
Alfred Naccash Avenue – Ashrafieh

Contents of the Issue

Global Journal of Computer Science and Technology: C
Software & Data Engineering
Volume 16 Issue 1 Version 1.0 Year 2016
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Efficient E-Government Services, Constraints and Problems of Technical Applications and Software and the Transition to Smart Government

By Dr. Yasser Elmalik Ahmed Seleman

Bisha University, Saudi Arabia

Abstract- That e-governance is the concept and structure of the system and the functions and activities of all activities and processes in e-business on the one hand the level of e-government and business on the other.

Because the government sector as a significant proportion of the total economic sectors in most countries of the world, and the fact that dealing with the public sector is not limited to the class and not others, but prevail all citizens and residents, institutions and others, and the fact that this multi-dealing in quality, methods and how it is done and models for different procedures and steps implemented and locations between the corridors of government departments, the concept of e-government came as an ideal way for the government to enable them to take care of the interests of the public from individuals and institutions electronically using cutting-edge technology without the need for the applicant to move between government departments.

Keywords: E:government is a modern introduction of the government of using the world wide web and the internet system in linking institutions to each other, and linking the various services private institutions and the general public, and the development of the information available to individuals in order to create a relationship specifications speed and accuracy designed to improve the quality of performance.

GJCST-C Classification : I.2.5 I.2.9

Strictly as per the compliance and regulations of:

Efficient E-Government Services, Constraints and Problems of Technical Applications and Software and the Transition to Smart Government

Dr. Yasser Elmalik Ahmed Seleman

Abstract- That e-governance is the concept and structure of the system and the functions and activities of all activities and processes in e-business on the one hand the level of e-government and business on the other.

Because the government sector as a significant proportion of the total economic sectors in most countries of the world, and the fact that dealing with the public sector is not limited to the class and not others, but prevail all citizens and residents, institutions and others, and the fact that this multi-dealing in quality, methods and how it is done and models for different procedures and steps implemented and locations between the corridors of government departments, the concept of e-government came as an ideal way for the government to enable them to take care of the interests of the public from individuals and institutions electronically using cutting-edge technology without the need for the applicant to move between government departments.

Keywords: E:government is a modern introduction of the government of using the world wide web and the internet system in linking institutions to each other, and linking the various services private institutions and the general public, and the development of the information available to individuals in order to create a relationship specifications speed and accuracy designed to improve the quality of performance.

I. Introduction

Scientific and technical developments in the era had to be institutions and centers of public and private information in various aspects of its activities and interests, to initiate the development of plans, and paint required for the development of information systems, policies, and communications, are linked to networks of national information, and the world, and interest in the preparation and the creation of technical personnel, specialized in the field of information , and networks, and communication systems [1].

Author: Kingdom of Saudi Arabia, Ministry of Education, Bisha University PhD in Computer Science, Omdurman Islamic University, M.sc in Information Technology, Newcastle (USA). M.sc in Information Technology, the National Ribat University.
e-mail: Dr.yaserking@hotmail.com

Researcher summarizes many of the e-government features and benefits and advantages in the following points .

That the application of e-government system provides many advantages are as follows:

1. performance of service speed: where it substituting Computer traditional manual system replaced, development occurred in the provision of service to the public where I said the time required to perform the service due to the flow of information and data from automated computer in connection with the required speed of service, and then be carried out at a time Set very short. This as well as the achievement of mail service is subject to the supervision easier and more accurate than those imposed on the employee in the performance of his works in the traditional management system.
2. Cost reduction: It is noted that the administrative performance of the business in the traditional way consumes very large amounts of papers and documents and stationery.

This as well as the need to offer more than one employee in order to review it and sign what is useful and to allocate it to another employee. It is the whole that would service high performance costs, due to the high prices and the prices necessary for the performance of service materials.

3. shortcut administrative procedures Do not doubt the traditional administrative work prevailing now characterized by a number of administrative complexities and because it needs most often to the approval of more than an administrative body on the work required, this as well as features that cause the employee-based performance of the service, which may take a vacation or It does not exist in the place of work and then hang the performance of the service from day to day.

In order to eliminate these bureaucratic follow it through e-government can simplify these procedures, and completed quickly and easily save time, effort and expenses and so especially with regard to places of departments and the number of workers.

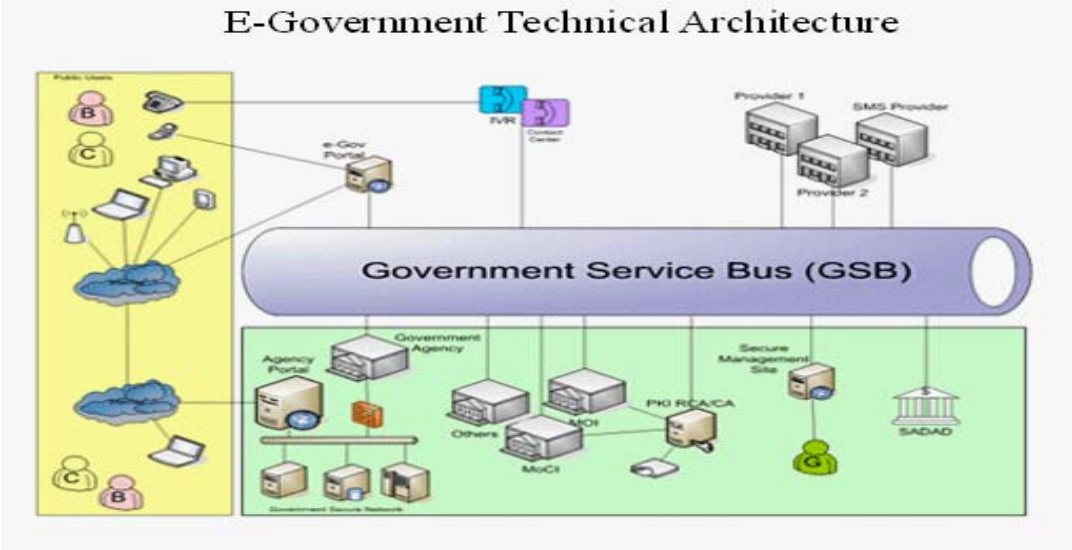

Fig. 1: Efficient E-government

II. Standards and Rules Governing own for the Production and Handling of Data and Adjust Circulated

It is as follows:

Technical requirements: It can be distributed into three main categories:

An infrastructure communications and Internet access requirements, and that can not be without them achieved any value at a broad level applications, also includes the provision of service centers, provide for citizens who do not have computer for their equipment to take advantage of the facilities offered by e-government.

IT infrastructure requirements, namely those related to the existence of an effective and capable of data from sources to collect information, and ensure the quality of this data and processed in accordance with the purposes of use systems. Requirements related tools software, including the availability of human frameworks eligible capable of dealing with these tools efficiently and effectively, it can be used for a broad spectrum of tools and software applications in order to achieve e-government. For example, there are: the rules and data banks, data and systems stores data mining, archiving and records management systems, Internet applications, geographic information systems, decision support, and other systems.

Those government applications usually are broad scope and magnitude of the volume of information they deal with, and this often requires a suitable combination of several tools and techniques to achieve the desired results.

III. Cultural Requirements

There is no doubt that the response of the citizens with renovations dictated by the transition to e-government will not be a direct, holiday of social behaviour and tendency to resist change are the key factors must be taken into account, if we want to experience e-government to succeed, so it is appropriate that the focus is initially on a simple, clear benefit applications and closely related to the daily lives of the people (which is what we find, for example, in the education and health sectors and transport), and so is associated with awareness and promote campaigns for citizens to deal with these applications, and the emphasis on the need for the interfaces of these applications by gravity and ease of use, and that is characterized by its content benefit and simplicity.

IV. Obstacles to the Application of E-Government

1. E-government application of the most important problems include information and communication technology in the postal sector deficit to meet customer requests on the phone, which are the most important channels of communication through online delivery.
2. The delay in completing the infrastructure for communications and variability from region to region, and building upon the great difference to bridge the digital worlds advanced countries in this field gap.
3. Limited deployment uses the Internet and that the proportion of users of this technology widespread globally is still weak.

We touched on a number of conclusions and recommendations which researcher deem necessary.

1. That e-government is the result of multiple shifts had information and communications technology frontrunner in. They perform the same traditional functions of government but through a unified electronic pattern.
2. The multiplicity of objectives through the application of this endeavor, which focused on reducing the administrative complexities, and strengthen transparency, speed, and ease of transactions.
3. The application of this project requires the provision of basic infrastructure for carrying out of a series of requirements, such as providing communications network, computer, Internet and the proliferation of specialized legislation in this area and the human element and other qualification.
4. One of the main obstacles to achieve in addition to the lack of infrastructure there is limited awareness of the work and awareness of the importance of this project and activation of knowledge management and failure to provide the necessary support prices on electronic gear and communications equipment and other. Smart government is the electronic services digital means for us dispense with many things, including the excessive use of paper and time lost in follow-up transactions between departments is an excellent step in the evolution of government services in the state system and the speed of completion of transactions and customer convenience in first class, which he could accomplish his business through his Smart phone without the need to go to the place of the government department and wait.

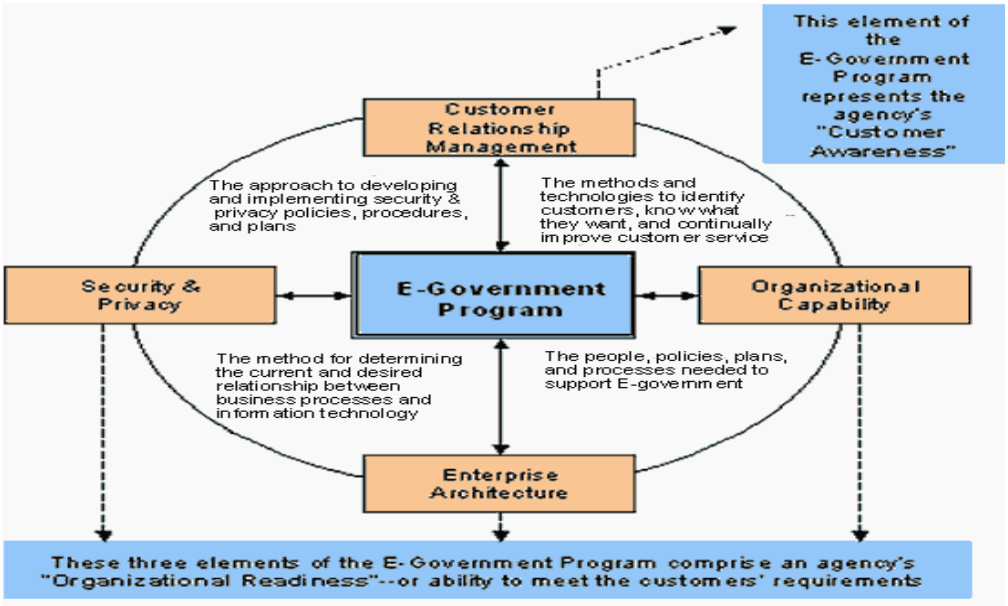

Fig. 2 : Smart government

a) Stages of the transition to smart government

Shift focus to smart government on the strategic use of the latest information and communication technologies, and on top of smart technology, to conduct a qualitative shift in the way they operate and accordingly government institutions, so as to achieve maximum user satisfaction, and effective cooperation with all relevant actors. This is done through the provision of means of communication seamless, interactive and intelligent work at any time and any place, across many devices.

b) Types of public improvements made smart government on E-government

it is necessary to understand the improvements that characterize the smart government to determine the quality of the services that must be taken into account in the development of smart services.

Comes with four different types of improvements made smart government on traditional way in which government agencies in providing services:

- Direct conversion of some of the e-government portal services: conversion of appropriate services between the existing services provided by the e-government portal to smart government services, which is a traditional electronic services are being provided on the smart platform.
- Clever new services are being provided to the public: a distinctive services may not be available in traditional e-government and made possible because of smart technologies. Examples of these services: payment of public transport and parking fees by using the mobile phone, in addition to the services that rely on geolocation.

- Services for field staff: means automating field force working; where are provided with government employees who work outside their offices (such as emergency and Inspection Services staff, and working to take care of patients at home) with a smart and techniques.

V. Recommendations

1. The need for effective investment in information and communication technology and provide the necessary infrastructure to build strong pillars of electronic government and that requires a spread of the Internet, and provide sponsors of this technology and the development of legislation and rehabilitation of the human element to ensure the overall generated by digital uses technical issues within cyberspace outstanding.
2. The need to sensitize and mobilize the citizens of the benefits and advantages of these technologies and provide the necessary facilities for them in connection with the acquisition of the necessary hardware costs, as occurred in Malaysia, Singapore, and other countries.

VI. Conclusions

In order for the e-government turn into a smart government will be working on several technical points and work smart government services on mobile phones and how they are assembled and endorsement serve individuals. Smart government may be provided through a government application of a uniform and deliberate government to disseminate guidance and general guidance on how to develop your technology has and how to design and contents of the service and how to protect service insurance (security and confidentiality of information) and then leave it for devices and various ministries in order to do internally government to develop Smart their own. Develop guidelines especially smart applications and templates (Smart Government Apps Guidelines).

Most governments has developed this special launch of a government Internet sites instructions but so far those governments did not work on the same application-level smart note that the time of citizen interaction with mobile apparatus far beyond the time consumed by that citizen interacting with browsers on desktop devices

References Références Referencias

1. United Nations Department of Economic and Social Affairs. "United Nations E-Government Survey 2014"(PDF). UN. Retrieved 2014-09-16.
2. Jump up^ OECD. The e-government imperative: main findings, Policy Brief, Public Affairs Division, Public Affairs and Communications Directorate, OECD, 2003
3. Jump up^ Grima-Izquierdo, C. (2010). A generic architecture for e-Government and E-Democracy: requirements, design and security risk analysis. Ed. LAP Publishing
4. China: The Next, Science Superpower, (2006.)
5. Essam Abdel Fattah rain, e-government between theory and practice, the new University House, Alozartih,2008 .
6. Abdel-Fattah Bayoumi Hijazi, e-government and legal system, the university thought Dar,..

Global Journal of Computer Science and Technology: C
Software & Data Engineering
Volume 16 Issue 1 Version 1.0 Year 2016
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Social Recommendation Algorithm Research based on Trust Influence

By Xue Yi, Hong Yinghan & Chen Pinghua

Guang Dong University of Technology, China

Abstract- Cold start and data sparsity greatly affect the recommendation quality of collaborative filtering. To solve these problems, social recommendation algorithms introduce the corresponding user trust information in social network, however, these algorithms typically utilize only adjacent trusted user information while ignoring the social network connectivity and the differences in the trust influence between indirect users, which leads to poor accuracy. For this deficiency, this paper proposes a social recommendation algorithm based on user influence strength. First of all, we get the user influence strength vector by iterative calculation on social network and then achieve a relatively complete user latent factor according to near-impact trusted user behavior. Depending on such a user influence vector, we integrate user-item rating matrix and the trust influence information. Experimental results show that it has a better prediction accuracy, compared to the state-of-art society recommendation algorithms.

Keywords : *collaborative filtering; cold start; data sparsity; social network; trust influence.*

GJCST-C Classification : *K.4.2 B.2.4*

Strictly as per the compliance and regulations of:

Social Recommendation Algorithm Research based on Trust Influence

Xue Yi [α], Hong Yinghan [σ] & Chen Pinghua [ρ]

Abstract- Cold start and data sparsity greatly affect the recommendation quality of collaborative filtering. To solve these problems, social recommendation algorithms introduce the corresponding user trust information in social network, however, these algorithms typically utilize only adjacent trusted user information while ignoring the social network connectivity and the differences in the trust influence between indirect users, which leads to poor accuracy. For this deficiency, this paper proposes a social recommendation algorithm based on user influence strength. First of all, we get the user influence strength vector by iterative calculation on social network and then achieve a relatively complete user latent factor according to near-impact trusted user behavior. Depending on such a user influence vector, we integrate user-item rating matrix and the trust influence information. Experimental results show that it has a better prediction accuracy, compared to the state-of-art society recommendation algorithms.

Keywords: collaborative filtering; cold start; data sparsity; social network; trust influence.

I. Introduction

As the third Internet revolution after the search engine, recommendation systems, without the need for exogenous information about either items or users, achieve the initiative to push personalized service by conducting analysis based on user past online behavior, such as ratings or usage. Collaborative filtering which is the classical recommendation algorithm, comprises memory and model based methods [1]. To solve the data sparsity and cold start conundrum, social recommendation has gradually become one of the important research fields in the recommender system. It usually introduces the corresponding user trust information in social network, and its underlying assumptions are: people tend to interact with the people who have similar interests and preferences, and in such process they would be more similar with each other [2].

Social network information can indeed be used to improve the accuracy of prediction score, but how to fully exploit the social network connection feature information becomes a hot topic among a number of researchers recent years. [3] proposed an approach (SoRec) based on probabilistic matrix factorization by assuming the user-specific and factor-specific latent feature follows Gaussian distribution, combined user trust information matrix T with user-item rating matrix R to constrain user latent feature vector, and then used the result to calculate rating matrix to predict; a combination of matrix factorization model and trust neighborhood-based model was proposed in [4] (RSTE); [5] presented an method that through active neighbors feature vectors to obtain the current user latent feature, and all user-specific matrix U, and then multiply the latent item-specific matrix V to forecast user-item ratings (Social MF).

These studies are mainly concentrated on the users directly connected each other in a social network, while ignoring the differences among user influence through out the network. [6] carried on a research from the perspective of trust and trusted, then combined each other with the addition (Trust MF). However, actually, trust and trusted information are not inseparable in a real a social network, and such trust information can be propagated, i.e. the user is not only influenced by neighbors, but also the users of greater influence in the network, in addition, during the stage that the trust and trusted models are mixed, since the trust and trusted detached from each other, tuning parameters becomes more difficult, resulting in poor interpretability even for a better prediction. Therefore, these algorithms above exist poor prediction accuracy issues.

In this paper we proposes a social recommendation algorithm based on the strength of trust influence: First of all, we conduct iterative calculation by using the connectivity of the trust network to obtain user-specific impact factors in the whole social trust network. Secondly, according to the influence differences among users, we constrain the user-specific latent factor matrix during the process of matrix decomposition, and by multiplying the user-specific and item-specific latent factor matrix to seek more accurate prediction scores.

II. Preliminaries

Conveniently, we describe the basic algorithm by the following example, while denote user User1 abbreviated as U1, as well Item1 as I1. Figure 1 is an abstraction for the real social network, in which each circle represents a user and the arrow points to the ones that can be trusted, e.g. user U3 trusted user U1, while

Author α σ ρ : School of Computer, Guang Dong University of Technology, Guangdong Guangzhou, 510006, China.
e-mail: andy.xueyi@gmail.com

the user U1 does not trust U3. We also use matrix to represent the trust information in the network, where 1 represents the trust, 0 mistrust, thus the matrix is non-symmetric matrix, denoted by T_n , n is the number of nodes, as shown in Table 1. User-item ratings information is shown as Table 2. We integrate the two datasets as input to get user-specific latent factor U and item-specific latent factor V by matrix decomposition, and then use the product of the two factors to get predicted ratings $\hat{R}$, as shown in table 3.We distinguish predicted ratings from known ones, by using the notation $\hat{R}$ for the predicted value of R .

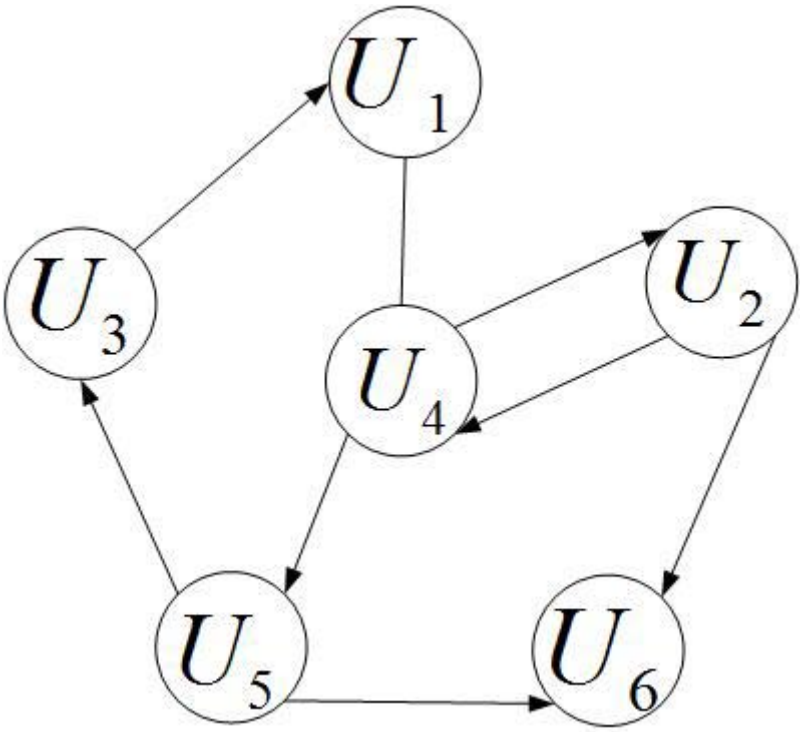

Fig. 1: Abstraction of Social Trust Network

Table 1: User trust information Matrix U

	U1	U2	U3	U4	U5	U6
U1	0	0	0	0	0	0
U2	0	0	0	1	0	1
U3	1	0	0	0	0	0
U4	1	1	0	0	1	0
U5	0	0	1	0	0	1
U6	0	0	0	0	0	0

Table 1 : user-item rating matrix V

	I1	I2	I3	I4	I5	I6
U1	5	2		3		4
U2	4	3			5	
U3	4		2			
U4						
U5	5	1	2		4	3
U6	4	3		2	4	

Table 2 : predicted user-item rating matrix $\hat{R}_{ij}$

	I1	I2	I3	I4	I5	I6
U1	5	2	2.5	3	4.8	4
U2	4	3	2.4	2.9	5	4.1

U3	4	1.7	2	3.2	3.9	3.0
U4	4.8	2.1	2.7	2.6	4.7	3.8
U5	5	1	2	3.4	4	3
U6	4	3	2.9	2	4	3.4

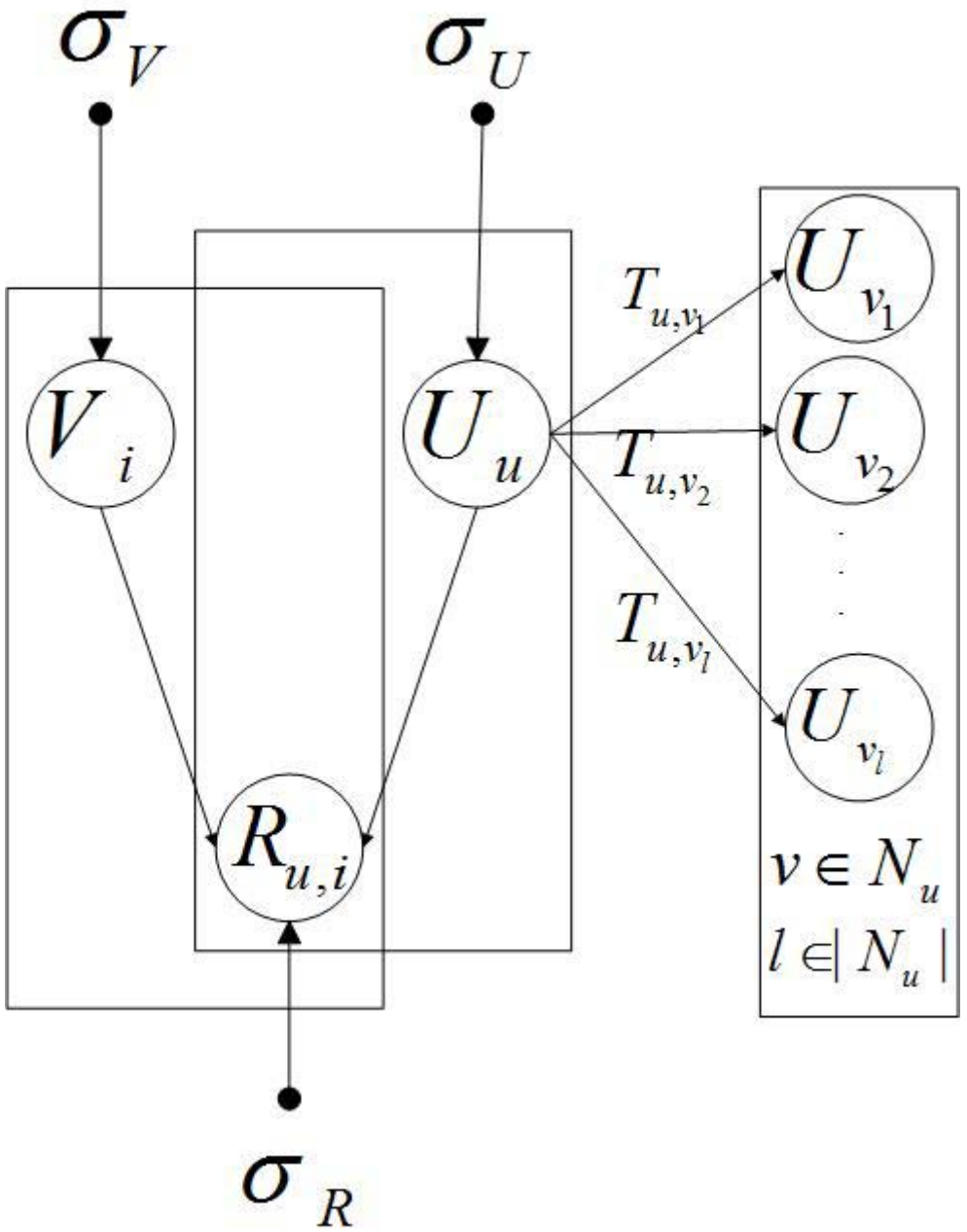

Fig. 2 : Factorization Diagram based on Rating Matrix and Social Network

According to the literature [7], introducing the social network into the user-rating ratings matrix can be treated as the rating matrix decomposition with a joint trust network matrix decomposition, as figure 2, they both constrain the user-specific latent factor U . We first define user-item rating matrix decomposition method, specifically address how to introduce a social network approach in the next section. We assume that the recommendation system involving m users and n items, the user-items rating matrix denoted as $R = [R_{ij}]_{mn}$, in which each rating R_{ij} indicates the preference by user i of item j ,where high values mean stronger preference, e.g. values can be integers ranging from 1 to 5. We denote an labeled matrix to show whether the user rates the item in user-item rating matrix

as $I=\left[I_{ij}\right]_{mn}$, if the user i has rated item j, set I_{ij} is 1, otherwise 0. By matrix factorization we can map the higher dimension matrix into a lower d dimension matrix. We denote user i latent factor as U_i, while item j latent factor as V_j, thus can get user-specific latent factor $U \in R^{d\times m}$, and item-specific latent factor $V \in R^{d\times n}$ and then use the product of user - specific and item-specific latent factor by U^TV to fit the available user-item ratings matrix R, whereby item not rated by users can be filled with a predicted score. The main learning function is obtained by minimizing the loss function as follows:

$$L=\sum_{i=1}^{m}\sum_{j=1}^{n}I_{ij}(U_i^TV_j-R_{ij})^2+\lambda_1\|U\|_F^2+\lambda_2\|V\|_F^2 \quad (1)$$

In the formula, we denote $\|\cdot\|_F^2$ as Frobenius norm, and in case of over-fitting, set parameters $\lambda_1,\lambda_2>0$ to control model complexity. In addition, a regularization method for weighting λ was proposed in [8] to avoid the parameters above in the model learning process, λ_1,λ_2 respectively account for user-specific and item-specific latent factor in the whole training model, at the same time, the introduction both of number of user U_i ratings and that of items rated is used to prevent the trained mode tilt to users who rated items too many or items rated by too many users, resulting in over-fitting problem, therefore, for the appropriate model items and users can be regularized as follows:

$$L=\sum_{i=1}^{m}\sum_{j=1}^{m}I_{ij}(U_i^TV_j-R_{ij})^2+\lambda_1\sum_i n_{u_i}\|U_i\|_F^2+\lambda_2\sum_j n_{v_j}\|V_j\|_F^2 \quad (2)$$

III. Algorithm Description

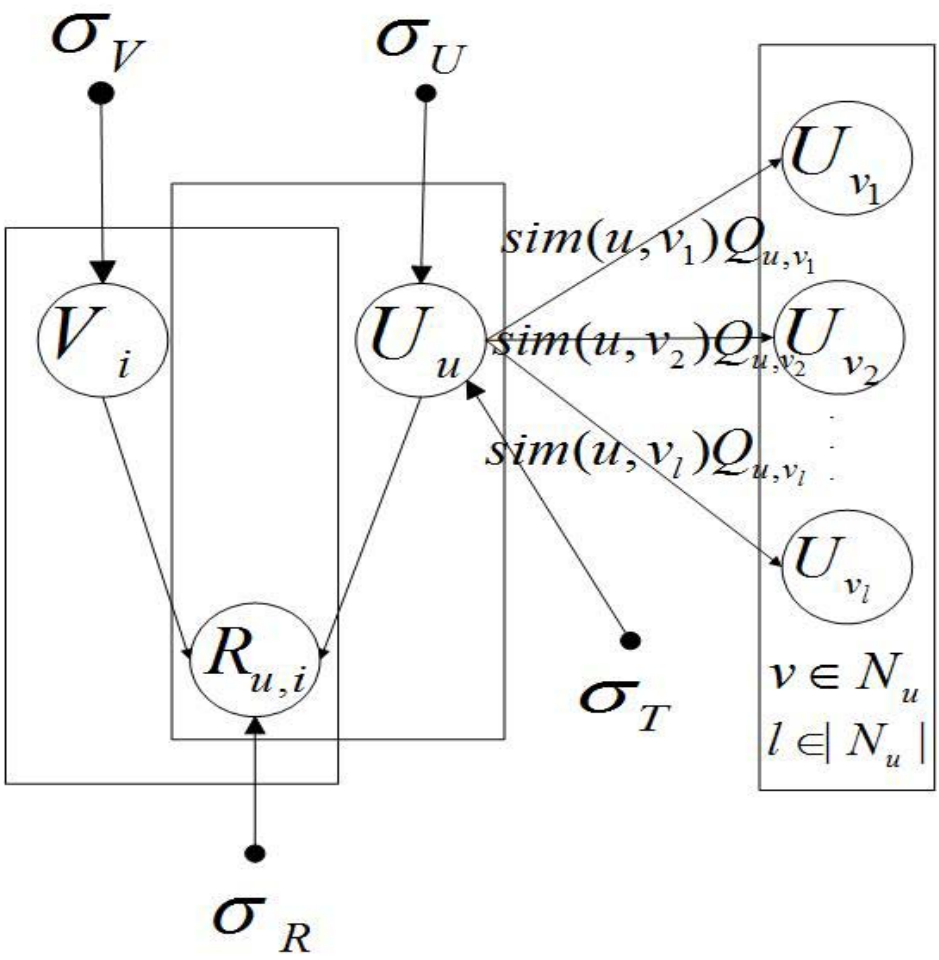

Fig. 3 : Our improved Algorithm T

The primary objective of this algorithm is to obtain the impact factors of user trust influence in the entire social network, which is utilized to measure the degree of users affected by their neighbors. We refer to the idea of Page Rank algorithm: the larger the number of user trusted throughout the trust network is, the bigger the trust influence such user has. We denote that there are m users involved in the ratings matrix R and trust matrix T, therefore, we can simultaneously both matrix decomposition operation by the same user-specific latent factor U. As mentioned in the section 1, for a given user trust network, denoted as $G := (M, N)$, which M represents user nodes, $T = \left[t_{m,n}\right]_{M,M}$ is denoted as adjacency labeled matrix, if user m trust user n, then $t_{m,n} = 1$, $t_{m,n} = 0$ and vice versa. Thus the user i trust influence can be defined as:

$$PR(i) = \frac{1-d}{M} + d \cdot \sum_{j \in T_i} \frac{PR(j)}{O_j} \tag{3}$$

M represents the number of whole nodes in the network, and $\Sigma_{i=0}^{N} \| PR(i) \| = 1$, T_i is denoted as user set that user i trusts, O_i is the number trusted by user i, and $d(0 \le d \le 1)$ is the damping factor, usually set to 0.85 [9]. Equation (3) only describes solution to achieve the influence of a particular user, so as to calculate the influence of the overall user-specific latent vector, it can be expressed as the formula (4):

$$E_{n+1} = dT_0 E_n + \frac{1-d}{N} I^+ \tag{4}$$

E_i represents the whole users trust influence in the i time iterative calculation, T_0 represents the original large-scale sparse trust matrix, I^+ is on behalf of a matrix filled with 1, and has the same dimension with T_0. When $n \to +\infty$, i.e. the matrix steps into the steady state:

$$E = dT_0 E + \frac{1-d}{N} I^+ \tag{5}$$

Whereby the equation (5) can be simplified as equation (6):

$$E = (I - dT_0)^{-1} \frac{1-d}{N} I^+ \tag{6}$$

When $0 < d < 1$, the solution exists. Using the iterative calculation described as equation (5), with the number increases, the model will gradually converge [14], then we can get entire user trust influence vector $E = [e_i]_{m \times 1}$, in the section above, we proposed a regularized method on social networking information. As for social network research, the majority of studies are based on the similarity of adjacent users, while ignore the fact that user trust someone does not mean they share a reliable presence of the same preference. If user trusted also has the same presence of target user behavior history information, the two are more likely to have same similar preferences, the objective function can be adjusted to the formula (7):

$$L = \sum_{i=1}^{m} \sum_{j=1}^{n} I_{ij} (U_i^T V_j - R_{ij})^2 + \lambda_1 \sum_i n_{u_i} \| U_i \|_F^2 + \lambda_2 \sum_j n_{v_j} \| V_j \|_F^2 + \lambda_3 \sum_{i=1}^{m} \sum_{j \in \Psi} Sim(i, j) \| U_i - U_j \|^2 \tag{7}$$

$Sim(i, j)$ is on behalf of cosine similarity between the user i and adjacent user j, Ψ is a labeled matrix indicating whether the user is directly connected to all the others in the matrix. Taking into account the diversity of user preferences, we thus make regularization to balance user similarity and user-specific latent factor, due to the fact that when two users are similar, $Sim(i, j)$ will become very large, and the user latent factor $\| U_i - U_j \|^2$ will become very small, and vice versa. In addition, users are more inclined to follow the similar users who are more influential and keep consistent in behavior with them, and that is the reason why we introduce the strength of trust influence, equation (8) is defined as below.

$$Q_{ij} = \frac{|e_j|}{\sum_{k \in \Psi} |e_{ik}|} \tag{8}$$

Q_{ij} indicates the trust proportion of user j in all user i directly connected users, we hope that user-specific latent factor tends to the overall average users, while taking into account the influence of adjacent users throughout the trust network, so we proposed a method which integrates both neighborhood similar user characteristics and their trust influence factors to associate a user-item ratings and trust network information, resulting in a more accurate model. In order to facilitate learning and weight adjustment, at first, we handle the data with normalization, as the trust value Q_{ij} is between 0 and 1, we normalize the original data by

using mapping function $f(x) = x / R_{max}$, so that its final value is between [0,1] . According to the literature [5], a logistic function $g(x) = 1/1 + exp(-x)$ was proposed to make the product of latent vectors to [0,1], so after the training, we can get the final predicted result by $g(U_i^T V_j) \cdot R_{\max}$. In view of keeping the same benchmarks among user-item ratings, social network data and the same constraint variable U, the new objective function is updated to equation (9):

$$L = \sum_{i=1}^{m}\sum_{j=1}^{n} I_{ij}(g(U_i^T V_j) - R_{ij})^2 + \lambda_1 \sum_i n_{u_i} \|U_i\|_F^2 + \lambda_2 \sum_j n_{v_j} \|V_j\|_F^2 + \lambda_3 \sum_{i=1}^{m}\sum_{j \in \Psi}^{n} Sim(i,j) \|U_i - Q_{ij} U_j\|^2 \quad (9)$$

During the training phase, we adopt the gradient descent method to minimize the objective function for solving the above-mentioned latent factors matrix, respectively U , V .

$$\frac{1}{2} \cdot \frac{\partial L}{\partial U_i} = \sum_{j=1}^{n} I_{ij} g'(U_i^T V_j)(U_i^T V_j - R_{ij}) V_j + \lambda_1 \sum_j n_{u_i} U_i + \lambda_3 \sum_{j \in \Psi} Sim(i,j)(U_i - Q_{ij} U_j) \quad (10)$$

$$\frac{1}{2} \cdot \frac{\partial L}{\partial V_j} = \sum_{i=1}^{m} I_{ij} g'(U_i^T V_j)(U_i^T V_j - R_{ij}) U_i + \lambda_2 \sum_i n_{v_j} V_j \quad (11)$$

When using Pearson correlation coefficients to conduct calculation of cosine similarity, since the scope of its value between $[-1,1]$, for taking a positive number, we use the mapping function $f(x) = (1+x)/2$ instead. $Sim(i, f)$ is the similarity value for user i and user f, R_{ij} represents the rating of user i for item j, $\overline{R_i}$ indicates user i average rating score, $I(i)$ is a labeled matrix indicating whether user i rate the item or not, as well as user f.

$$Sim(i, f) = \frac{\sum_{j \in I(i) \cap I(f)} (R_{ij} - \overline{R_i}) \cdot (R_{fj} - \overline{R_f})}{\sqrt{\sum_{j \in I(i) \cap I(f)} (R_{ij} - \overline{R_i})^2} \sqrt{\sum_{j \in I(i) \cap I(f)} (R_{fj} - \overline{R_f})^2}} \quad (12)$$

IV. Complexity Analysis

The model training time consists two parts, the first part is the algorithm calculate the user trust influential feature vector, the second parts is for solving the latent factors by using the gradient descent method. We conduct iterative calculation to get the trust influential feature vector, if the entire number of nodes in the trust network is m, the average number of users trusted by per user is n, usually n is a relatively small number. Given t_1 times of iterations to achieve convergence, time complexity is $O(mnt_1)$, when the data is in large scale, the literature [10] proposed a distributed solution that can significantly reduce the time.

The time complexity of solving L is $O(t_2 d(|\Omega| + N))$. t_2 is a specified number of iterations, d is the dimension of latent factors, $|\Omega|$ is the scale of the observed ratings, N represents time complexity of similarity calculation, since the matrix is so sparse that the time complexity $O(N) << O(n^2)$. After the combination of these two steps, the total time complexity of the algorithm model is $O(t_2 d(|\Omega| + N) + mnt_1)$.

V. Experiments and Results

a) Dataset Selection

The classical data set were Epinions (665KB) [11], which contains not only the user-item ratings data, but also trust relationship between users, and thus it is usually recommended as baseline data set to test social recommendation algorithms. We use it for testing and validating the currently mainstream recommendation algorithms as well as our improved algorithm. First of all, we made the basic statistics of the data set as shown in Table 4. Trust (degree) and trusted (out-degree) statistics information shows that when the trust or be trusted number of persons increases, the corresponding statistics number gradually reduces, which follows the power law distribution.

b) Cross-Validation

We use 5-fold cross-validation methods for training and testing the models. For each test we randomly selected 80% of the whole data as the training data set and the remaining 20% is used for testing. The next experiment results discussed in the final comparison is obtained by averaging the results of tests from five repeated times.

c) Evaluation

Evaluation criteria used in the Experiments are based on the average absolute error MAE and root mean square error RMSE:

$$MAE = \frac{\Sigma_{i,j} |R_{ij} - \hat{R}_{ij}|}{N} \quad (13)$$

$$RMSE = \frac{\sqrt{\Sigma_{i,j} | R_{ij} - \hat{R}_{ij} |^2}}{N} \quad (14)$$

Table 4 : Basic statistics of the data set Epinions

Basic Data					Scenes			Others	
Users	Items	Rating Scale		Density	Users	Link Types		Items	tags
40,163	139,738	664,824	[1,5]	0.000118458	66,807	487,183	Trust	General	None

Table 6 : Results in Dimension 5 and 10

	Evaluation	SoRec	RSTE	SocialMF	TrustMF	T
d=5	MAE	0.9197	0.8635	0.8826	0.8212	0.8010
	RMSE	1.151	1.1071	1.1107	1.0585	1.0320
d=10	MAE	0.9152	0.8572	0.8567	0.8148	0.8059
	RMSE	1.1773	1.1483	1.1113	1.0771	1.0408

From the table, we can conclude that, regardless of index value of MAE or RMSE, the smaller its value is, the more accurately the algorithm predicts the score.

d) Comparison

In consideration of the similar algorithms, our main work is integrating users trust influence in the social network into the user-item ratings. Thus while making the lateral comparisons, in addition to the original matrix decomposition algorithms, we also made comparisons among social recommendation algorithms which are focus on trust.

e) Contrast Algorithms and Reasons

When training the models, we set parameters of each algorithms by referring to the literature [8] [9], which introduced the optimal parameters of each algorithms. Our algorithm is denoted as T and more details described as the following table 5:

Table 5 : Contrast Algorithm and Params Settings

Algorithm Name	Parameters
SoRec	$\lambda_U = \lambda_V = \lambda_Z = 0.001, \lambda_C = 1$
RSTE	$\lambda_U = \lambda_V = 0.001, \alpha = 0.4$
SocialMF	$\lambda_U = \lambda_V = 0.001, \lambda_T = 1$
TrustMF	$\lambda = 0.001, \lambda_T = 1$
T	$\lambda_1 = \lambda_2 = 0.001, \lambda_3 = 1$

For all the algorithms above based on matrix decomposition, the dimension of latent factors are set to 5 and 10 respectively, and during model training stage，the same initialized strategy is adopted: the original matrixes involved are filled with a random number uniformly distributed values between 0 and 1.

f) Results

Experiment results show that with the increase of the number of iterations, the index value of RMSE and MAE keeps declining, which indicates the introduction of the trust influence factors has a positive impact on enhancing the overall effectiveness of recommendation and improves the accuracy of predicted results. On the dataset, we have achieved results as showed in Table 6, while our algorithm is denoted as T. Cross-validation results show the performance of the our proposed algorithm is slightly better than all the other compared algorithms. In addition, we conducted statistical analysis of RMSE and MAE index value during the iterative calculation and found that with increasing number of iterations, RMSE and MAE index value gradually reduce and eventually keep stabilized, indicating that the algorithm utilizes influence characteristics among users in the trust network to predict user rating propensity is relatively effective and eventually enhances the accuracy of the prediction score of recommendation.

VI. Conclusion

As the common recommendation system algorithm, collaborative filtering encountered sparse data and cold start problems, which leads to poor accuracy. Given the fact that the impact of opinion leaders for individuals in the social network, this paper proposed one social recommendation algorithm based on trust influence to alleviate such problems above. Compared to the previous research merely on adjacent user preferences, we consider the trust propagation mechanism and try out iterative calculations based the connectivity of social trust network to obtain user influence values, and then integrate both similar neighbor characteristics and their trust influence factors. Experiments show that compared the most state-of-art social algorithms, our approach optimizes the recommended results and improves the prediction accuracy. However, obviously, this algorithm complexity positively correlates with the number of iterations, when

faced with large-scale data, the computation time is too long, especially for single node. Therefore, the next research focus on the implementation of distributed algorithm for the iterative calculations involving large-scale data.

References Références Referencias

1. Su, X., & Khoshgoftaar, T. M. (2009). A survey of collaborative filtering techniques. Advances in Artificial Intelligence, 2009(12).
2. Kautz, H., Selman, B., & Shah, M. (1997). Referral web: combining social networks and collaborative filtering. Communications of the Acm, 40(3), 63-65.
3. Ma, H., Yang, H., Lyu, M. R., & King, I. (2008). So Rec: Social Recommendation Using Probabilistic Matrix Factorization. International Conference on Information & Knowledge Management (pp.931-940). ACM.
4. Ma, H., King, I., & Lyu, M. R. (2009). Learning to recommend with social trust ensemble. Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval (pp.203-210). ACM.
5. Jamali M, Ester M.(2010). A matrix factorization technique with trust propagation for recommendation in social networks. Proceedings of the fourth ACM conference on Recommender systems (pp.135-142). ACM.
6. Yang, B., Lei, Y., Liu, D., & Liu, J. (2013). Social collaborative filtering by trust. Proceedings of the Twenty-Third international joint conference on Artificial Intelligence (pp.2747-2753). AAAI Press.
7. Salakhutdinov, R., & Mnih, A. (2012). Probabilistic matrix factorization. Nips, 1257-1264.
8. Ma, H., Zhou, D., Liu, C., Lyu, M. R., & King, I. (2011). Recommender systems with social regularization. Fourth Acm International Conference on Web Search & Data Mining (pp.287-296). ACM.
9. Amy N. Langville, & Carl D. Meyer. (2003). Deeper inside pagerank. Internet Mathematics, 1(3), 335-380.
10. Sarma, A. D., Molla, A. R., Pandurangan, G., & Upfal, E. (2013). Fast Distributed Page Rank Computation. Distributed Computing and Networking. Springer Berlin Heidelberg.
11. Epinions data set: http://www.trust .org/wiki/Epinions

Global Journal of Computer Science and Technology: C
Software & Data Engineering
Volume 16 Issue 1 Version 1.0 Year 2016
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

On Database Relationships Versus Mathematical Relations

By Christian Mancas

Bucharest Polytechnic University, Romania

Abstract- Unfortunately, the widespread used one-to-many, many-to-one, one-to-one, and many-to-many database relationships lack precision and are very often leading to confusions that affect the quality of conceptual data modeling and database design. This paper advocates replacing them with the rigorous math notions of relations and (one-to-one) functions.

GJCST-C Classification : *F.4.1 H.2.3*

Strictly as per the compliance and regulations of:

On Database Relationships versus Mathematical Relations

Christian Mancas

Abstract- **Unfortunately, the widespread used one-to-many, many-to-one, one-to-one, and many-to-many database relationships lack precision and are very often leading to confusions that affect the quality of conceptual data modeling and database design. This paper advocates replacing them with the rigorous math notions of relations and (one-to-one) functions.**

I. Introduction

The widely used Entity-Relationship (E-R) Data Model (E-RDM, e.g. [Chen, 1976], [Thalheim, 2000], [Mancas, 2015]) is and will continue to be successful in database (db) design mainly due to the graphical nature of its E-R Diagrams (E-RDs) and simplicity.

a) E-RDs

In its original version [Chen, 1976], atomic (*entity*-type) object sets are represented in E-RDs by rectangles, compound (*relationship*-type) ones by diamonds, and the Relational Data Model (RDM, e.g.[Codd, 1970], [Abiteboul et al., 1995], [Mancas, 2015]) *attributes* (object set properties) by ellipsis (attached to the corresponding rectangles and diamonds).

Structural E-RDs only contain rectangles and diamonds (which connect rectangles), without any ellipsis. As such, they are non-directed graphs whose nodes are rectangles and diamonds and whose edges are so-called "*roles*" (of the connected entity-type object sets in the corresponding relationship-type ones).

Figure 1 shows an example of a Chen-style E-RD, while Figure 2 presents the corresponding structural one. Obviously, *CITIES* and *COUNTRIES* are entity-type object sets, *CITIES_COUNTRIES* and *COUNTRIES_CAPITALS* are relationship-type ones, *belongs to*, *has*, *is capital*, and *has capital* are roles, whereas *Name*, *Zip Code*, *Population*, *Code*, *Tel Prefix* are attributes.

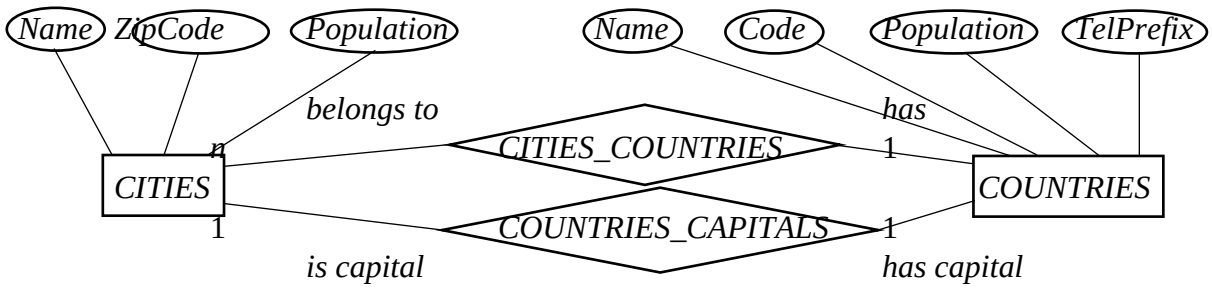

Figure 1 : An example of a Chen-style E-RD

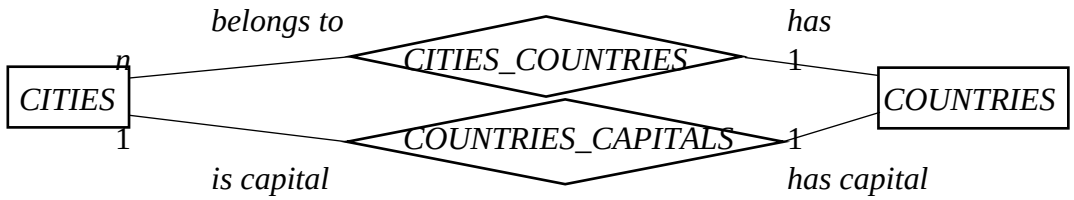

Figure 2 : The Chen-style structural E-RD corresponding to the one of Figure 1

Author: Bucharest Polytechnic University.
e-mail: christian.mancas@gmail.com

Roles have associated cardinalities. For example, read from left to right, *CITIES_COUNTRIES* is said to be a *many-to-one relationship* (as there generally are many cities in a country) and this is why *belongs to* has cardinality *n*, while *has*has 1. Obviously, read from right to left, it is a *one-to-many* relationship (as generally a country has many cities). Similarly, *COUNTRIES_CAPITALS* is said to be a *one-to-one relationship* (as countries may have only one capital and any city may be the capital of only one country) and this is why both *is capital* and *has capital* have cardinality 1.

Figure 3 shows a so-called *many-to-many relationship* (as any person may get married several times with different persons), where both roles have cardinality *n*.

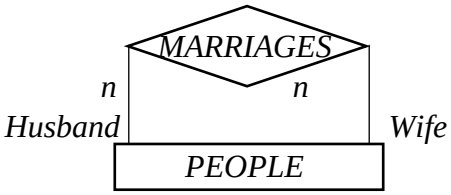

Figure 3 : An example of a many-to-many relationship

We are using a slightly different notation [Mancas, 2015]: just like in its original version, atomic (entity-type) object sets are represented by rectangles, mathematic non-functional relation type ones (i.e. subsets of Cartesian products) are represented by diamonds, but functional ones are represented as arrows, just like in math. Hence, in our version structural E-RD (from now on abbreviated as E-RD) are oriented graphs whose nodes are only object sets and whose edges are *structuralfunctions* (i.e. functions defined on and taking values from object sets[1]).

For example, as, in fact, both *CITIES_COUNTRIES* and *COUNTRIES_CAPITALS* are functional, Figure 4 shows the equivalent of the Chen-style E-RD from Figure 2.

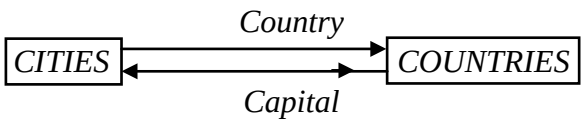

Figure 4 : The math-style E-RD equivalent to the one in Figure 2

As *MARRIAGES* is not functional, our math-type notation is identical to the Chen-type one from Figure 3.

b) Corresponding mathematical relations

Recall that, algebraically, a relation is a non-empty subset of a Cartesian product. First (minor) difference of db relationships as compared to math relations is that they may be empty (at least immediately after they are declared and up to the moment when a first element is inserted into their instances, but possibly also afterwards, whenever their instances are emptied by deleting all of their elements and up to the moment when new elements are again inserted into them).

Second (major) difference between them is that the math ones are positional (as Cartesian products are non-commutative), whereas db ones are not: they only require that all roles of any relationship be pairwise distinct.

For example, mathematically, $CITIES \times COUNTRIES \neq COUNTRIES \times CITIES$, which means that when both relationships from Figures 2 and 5 are read either from left to right or from right to left they are distinct, whereas from the db perspective they are strictly equivalent, no matter how are they read (which would correspond to the equivalence classes of Cartesian products immune to the permutations of their member sets).

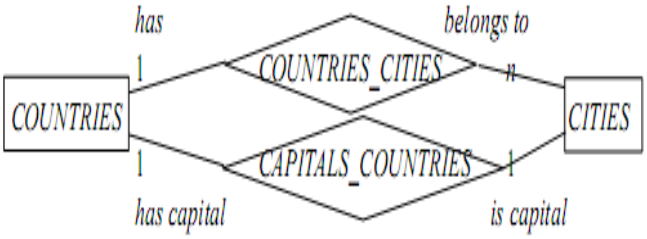

Figure 5 : The E-RD dual to the one of Figure 2

Another advantage of our notation (beside simplicity and math compatibility) becomes clear when comparing Figure 2 with its corresponding dual from Figure 6: no relationship-type set name has to change – only arrow directions are reversed.

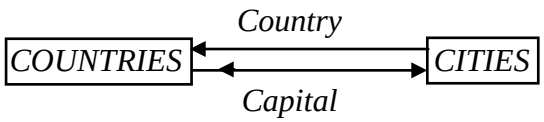

Figure 6 : The E-RD dual to the one in Figure 4

Also recall that there is a very important particular case of math relations, namely the functions (mappings); a *function* is a binary relation satisfying two additional constraints: it is totally defined and it is functional. Read from left to right, the first set is called the *domain*, while the second is called *co-domain*. For example, the function *Country* :*CITIES*→*COUNTRIES* has domain *CITIES* and co-domain *COUNTRIES* and it is a function because is *totally defined* (that is any city belongs to a country) and *functional* (i.e. any city belongs to only one country).

Database functions (which in relational ones are implemented as table columns) differ slightly from math ones only because totality is not compulsory: for

[1] as compared to attribute-type ones, also defined on object sets, but taking values into (subsets of) data types (e.g. Population : CITIES→ [0, 3*10^6])

example, as capitals might not be temporarily known or of interest for any country, the function *Capital* :*COUNTRIES*→*CITIES*may not be totally defined.

Totality is considered in dbs as a constraint that has to be explicitly asserted whenever desired. For example, in the (Elementary) Mathematical Data Model ((E)MDM, e.g. [Mancas, 1990], [Mancas, 2016]), the complete declaration of *Country* is *Country* : *CITIES*→*COUNTRIES*, *total*. In RDM, this is called a *not-null constraint*, meaning that the corresponding column does not accept *null values* (i.e. distinguished values represented either as null strings or with the keyword <NULL>). Considering a countable distinguished set NULLS, a possible dual (E)MDMnotation for the above two functions is *Country* : *CITIES*→*COUNTRIES* and *Capital* : *COUNTRIES*→*CITIES*∪ NULLS, respectively, in which case total definition is always satisfied, just like in math.

Obviously, *Capital* is a one-to-one function, i.e. one for which to any pair of distinct domain elements corresponds a pair of distinct function values. This is why, in our notation (e.g. Figures 4 and 6) its arrow is a double one, and its complete (E)MD Mdefinition is *Capital* : *COUNTRIES*↔*CITIES*.

Note that roles of non-functional relationships (e.g. *Husband* and *Wife* from Figure 3 above) are also structural functions, namely canonical Cartesian projections (e.g. *Husband* : *MARRIAGES* →*PEOPLE*, *Wife* : *MARRIAGES* →*PEOPLE*).

II. Disadvantages of using DB Relationships Instead of Math Relations and Functions

There is only one advantage in using E-RD relationships, especially when using our simpler and math-type notation: the fact that they are graphic (and a good picture is worth thousand words). Unfortunately, there are much more important disadvantages as well.

a) Unnaturalness of Chen-type functional relationships

Representing functional relationships as diamonds has several pitfalls:

- It is true that, being particular cases of binary relations, they can be thought of as object sets as well (in particular, sets of elements of the type <*x*, *f*(*x*)>), but, in fact, both mathematically (which considers them functions, not sets) and from the db point of view (which, by applying the Key Propagation Principle [Mancas, 2015], implements them as table columns, in particular foreign keys) they are not dealt with as such, just like the non-functional ones (which are implemented as tables, just like for the entity-type ones).
- Their names are confusing: obviously, for example, both *Country* and *Capital* are much clearer than *CITIES_COUNTRIES* and *COUNTRIES_CAPITALS*; a clear sign that they are unnatural objects is that they lack natural names, which only exist for non-functional relationships (e.g. *STOCKS* instead of *WAREHOUSES_PRODUCTS*).
- The need for three distinct names (for the relationship and its two roles) instead of only one (the function) is also unnatural. Again, as compared to non-functional relationship role names, which are natural (e.g. *Husband*, *Wife*, *Product*, *Warehouse*, *Home Team*, *Visiting Team*, etc.), they generally have an Artificial Intelligence flavor (e.g. *is*, *has*, *belongs*, etc.), not a db or math one.
- The redundancy of one-to-many relationships: as we read math from left to right, functions are many-to-one relationships; one-to-many ones are the same corresponding functions, but read from right to left (i.e. from the co-domain to the domain).

b) Confusion between one-to-oneness and bijectivity

Not only beginners, but also, for example, MS *Access* designers are confusing one-to-oneness with bijectivity. For example, if you first declare *Capital* as a (unique) key (i.e. as being one-to-one) and then try to enforce its referential integrity, depending on the instances of the two tables it relates, you might not succeed in either enforcing it (when there are more cities than countries, whichis the norm) or inserting data in any of the two involved tables (when both instances are empty, enforcing referential integrity succeeds, but then you may not enter either cities, as there are no corresponding countries, or countries, as there are no corresponding cities).

This is clearly due to the confusion done between one-to-oneness and bijectivity (i.e. one-to-oneness and ontoness).[2]

c) The many-to-many relationships trap

The worst issue with db relationships is that they may not even correspond to object sets.

For example, if you enforce uniqueness of elements in the above *MARRIAGES* (i.e. uniqueness of the product *Husband*•*Wife*), then you may not store re-marriages (e.g. Elisabeth Taylor and Richard Burton married and divorced each other several times). If you do not enforce it, then it is not even a set, as it accepts duplicates.

Generally, you have to validate data modeling correctness for each relationship, by checking the one-to-oneness of the product of all of its roles: if it is not (like for *MARRIAGES*, where *Husband*•*Wife*is not one-to-one), then the corresponding relationship is ill-defined (and either it lacks at least another role or it is, in fact an entity-type object set).

Consequently, the correctmodel in all contexts in which divorce (hence, remarrying) is possible is the one in Figure 7:

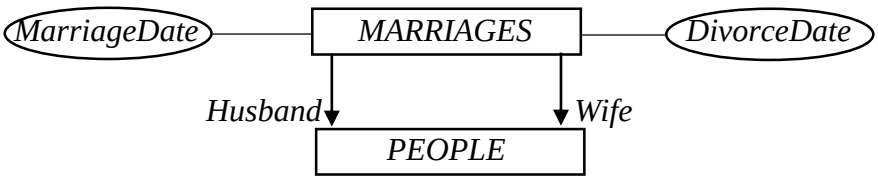

Figure 7 : Correct data model of MARRIAGES (as an entity, not relationship-type object set, like it is incorrectly modeled in Figure 3)

III Conclusion

To conclude with, during conceptual data modeling and db design it is always much, much better to think in terms of math relations and functions, rather than in those of one-to-many, many-to-one, one-to-one, and many-to-many ones.[3]

Otherwise, you risk confusions between one-to-many and many-to-one ones, one-to-oneness and bijectivity, and even between relationship and entity-type object sets.

Moreover, our E-RD notations [Mancas, 2015] are much simpler, natural, and close to math than the original ones.

References Références Referencias

1. [Abiteboul et all, 1995] Abiteboul, S., Hull, R., Vianu, V. (1995)*Foundations of Databases*.Addison-Wesley: Reading, MA.
2. [Chen, 1976] Chen, P. P. (1976). The entity-relationship model: Toward a unified view of data. *ACM TODS* 1(1), 9-36.
3. [Codd, 1970] Codd, E. F. (1970). A relational model for large shared data banks. CACM 13(6), 377-387.
4. [Mancas, 1990] Mancas, C. (1990). A deeper insight into the Mathematical Data Model. 13th Intl. Seminar on DBMS, ISDBMS, Mamaia, Romania, 122–134.
5. [Mancas, 2015]Mancas, C. (2015). *Conceptual Data Modeling and DB Design. A Fully Algorithmic Approach. Vol. I: The Shortest Available Path*, Apple Academic Press, NJ.
6. [Thalheim, 2000] Thalheim, B. (2000) *Fundamentals of Entity-Relationship Modeling*. Springer-Verlag, Berlin.

[2] Fortunately, there is a workaround for it in MS *Access* too: if you first enforce referential integrity and only then uniqueness, no issue arises.

Global Journal of Computer Science and Technology: C
Software & Data Engineering
Volume 16 Issue 1 Version 1.0 Year 2016
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Cluster Analysis of Medical Research Data using R

By Lavanya Pamulaparty, Dr. C.V Guru Rao & Dr. M. Sreenivasa Rao

JNT University, India

Abstract- Cluster analysis divides the data into groups that are meaningful, useful or both. It is also used as a starting point for other purposes of data summarization. This paper discuss some very basic algorithms like K-means, Fuzzy C-means, Hierarchical clustering to come up with clusters, and use R data mining tool. The results are tested on the datasets namely Online News Popularity, Iris Data Set and from UCI data repository and mi RNA dataset for medical data analysis. All datasets was analyzed with different clustering algorithms and the figures we are showing is the working of them in R data mining tool. Every algorithm has its uniqueness and antithetical behavior.

Keywords: *k-means algorithm, fuzzy c-means algorithm, hierarchical clustering algorithm, r tool.*

GJCST-C Classification : *H.3.3 I.5.3*

Strictly as per the compliance and regulations of:

Cluster Analysis of Medical Research Data using R

Lavanya Pamulaparty [α], Dr. C.V Guru Rao [σ] & Dr. M. Sreenivasa Rao [ρ]

Abstract- Cluster analysis divides the data into groups that are meaningful, useful or both. It is also used as a starting point for other purposes of data summarization. This paper discuss some very basic algorithms like K-means, Fuzzy C-means, Hierarchical clustering to come up with clusters, and use R data mining tool. The results are tested on the datasets namely Online News Popularity, Iris Data Set and from UCI data repository and mi RNA dataset for medical data analysis. All datasets was analyzed with different clustering algorithms and the figures we are showing is the working of them in R data mining tool. Every algorithm has its uniqueness and antithetical behavior.

Keywords: k-means algorithm, fuzzy c-means algorithm, hierarchical clustering algorithm, R tool.

I. Introduction

Cluster analysis divides data into meaning full groups (clusters) which share common characteristics i.e. same cluster are similar to each other than those in other clusters. It is the study of automatically finding classes. A web page especially news articles which are flooded in the internet have to be grouped. The clustering of these different groups is a step forward towards the automation process, which requires many fields, including web search engines, web robots and data analysis.

Any new web page goes through numerous phases including data acquisition, preprocessing, Feature extraction, classification and post processing into the database. Cluster analysis can be regarded as a form of the classification which creates a labeling of objects with class labels. However it derives these labels only from the data. Data mining functionalities are the Characterization and discrimination, mining frequent patterns, association, correlation, classification and prediction, cluster analysis, outlier analysis and evolution analysis [1].

Clustering is a vivid method. The solution is not exclusive and it firmly depends upon the analysts' choices. Clustering always provides groups or clusters, even if there is no predefined structure. While applying cluster analysis we are contemplating that the groups exist. But this speculation may be false. The outcome of clustering should never be generalized. [9].

Author α : Department of CSE, Methodist college of Engg. & Tech., OU, Hyderabad. e-mail: lavanya.post@gmail.com
Author σ : Department of CSE, S R Engineering College, JNT University, Warangal.
Author ρ : Department of CSE, School of IT, JNT University, Hyderabad.

II. R Tool

R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and Mac OS [12]

R is public domain software primarily used for statistical analysis and graphic techniques [10]. A core set of packages is included with the installation of R, with more than 7,801 additional packages (as of January 2016[update]) available at the Comprehensive R Archive Network (CRAN), Bio conductor, Omegahat, Git Hub, and other repositories.[14] R tool provides a wide class of statistical that includes classical statistical tests, linear and nonlinear modeling, classification, time-series analysis, clustering and various graphical functions.[13]

R uses collections of packages to perform different functions [11]. CRAN project views provide numerous packages to different users according to their taste. R package contain different functions for data mining approaches. This paper compares various clustering algorithms on datasets using R which will be useful for researchers working on medical data and biological data as well. For this IDE, R Studio is used refer the below Figure 1.

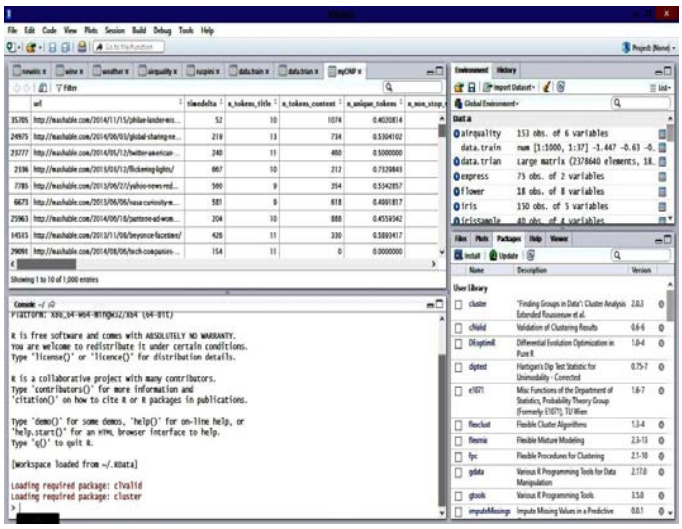

Fig 1: R tool Studio

III. Clustering Algorithms

a) K-Means

The term "k-means" was first used by James Macqueen in 1967 [2], though the idea goes back to

1957 [4]. The K-means algorithm is the most commonly used and simplest method among all partitional clustering algorithms. As Harington and Wong1979 mentioned it is an iterative method which minimizes the sum of the squares for a given number of Clusters.

Here's how the algorithm works [5]:

1. Select k points as initial centroids.
2. Repeat
3. From K clusters by assigning each point to its closest centroid.
4. Recompute the centroid of each cluster.
5. Until centroids do not change.

K-Means reaches a state in which no points are shifting from one cluster to another e.g. repeating until only 1% of the points change clusters.

For measuring the quality of the clustering we measure Sum of the squared error (SSE) or scatter.

$$SSE=\sum_{i=1}^{k}\sum_{x\in c_i} dist(c_i , x)^2$$

Where dist is standard Euclidean distance between two objects in Euclidean space. The centroid (mean) of the ith cluster that minimizes the SSE is defined as

$$\sum_{i=1}^{k}\sum_{x\in c_i} dist(c_i , x)^2$$

The advantage of this method is highly scalable of the huge sum of data sets with O(n * k * r) where r is the number of rounds, where n represent number of data items, k represent numbers of clusters [14]. It has user defined constant K and Runtime is totally dependent on the initial pick of centroids.

b) K-Means Implementation using R

For this analysis we have considered Online News Popularity datasets which consists of articles published by Mash able (www.mashable.com) [4]. Instances are 39797 and Number of Attributes is 61. As the results of the k means are undeterministic, we have followed the practice of running multiple rounds of k means so performed on various k values as k=3, k=5 and k=10. The best iteration is one who minimizes the average distance of each point to its assigned centroid.

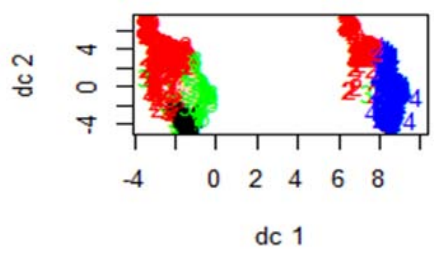

Fig 2 : K means plot with k=5

The Figure 2 shows the results of clustering the Online News Popularity datasets with 39645 number of News articles URL's. The above results show that there is overlapping of clusters. So preprocessing is required to address this problem and omit the NA values also. Then the following is the code after data cleaning [15].

```
>onpdat<- kmeans (myONP [, c ("n_tokens_title","n_tokens_content")], centers=5, nstart=10)
> Onpdata
        K-means clustering with 5 clusters of sizes 15, 67, 195, 319, and 404
Cluster means:
  n_tokens_title n_tokens_content
1       11.46667        2648.9333
2       10.29851        1469.2985
3       10.41026         897.5897
4       10.50470         492.4044
5       10.28713         211.3886
Clustering vector:
35705 24975 23777  2336  7785  6673 25963 14515
29091 18807 27116 37480 14360 29375 35316  8015
24621  4744 10096 14587
    3     3     4     5     4     4     3     5     5     5     5     4     5
3     5     4     4     5     4     3
 506 14445 32852 12857 18210 22647 18642  7034
31249 25246 29996  4077 27331 15531 31001 24434
29564 20883 20002 29804
    3     4     4     4     5     3     4     5     5     5     3     5     4
3     5     5     5     5     2     5
 881 18018 27648 26153 18032 32512 16539  9241
2668  3755 11938 19576 24987 15355 34454 11081
17326 12545 24563  9737
    4     2     4     3     5     3     3     5     5     2     5     5     5
4     5     5     4     5     3     1
Within cluster sum of squares by cluster:
[1] 3437391 3417672 3385646 3069653 3279165
 (between_SS / total_SS =  92.1 %)
Available components
[1] "cluster"      "centers"      "totss"        "withinss"
"tot.withinss" "betweenss"    "size"
[8] "iter"         "ifault"
> summary(onpdata)
             Length  Class   Mode
cluster      1000    -none-  numeric
centers        10    -none-  numeric
totss           1    -none-  numeric
withinss        5    -none-  numeric
tot.withinss    1    -none-  numeric
betweenss       1    -none-  numeric
size            5    -none-  numeric
iter            1    -none-  numeric
ifault          1    -none-  numeric
library(MASS)
> parcoord(data.train,onpdata$cluster)
>confuseTable.km<-
table(myONP$n  tokens  title,onpdata$cluster)
> confuseTable.km
     1  2  3  4  5
 3   0  0  0  0  1
```

```
5   0  0  0  1  2
6   0  1  1  8 13
7   1  8 10  7 26
8   1  4 29 34 44
9   2 12 25 48 53
10  2 12 39 67 81
11  1 10 35 55 77
12  2  7 22 53 50
13  4  9 24 24 19
14  1  3  7 12 25
15  0  1  2  6  9
16  0  0  0  4  3
17  0  0  0  0  1
18  1  0  1  0  0
> library(flexclust)
> randIndex(confuseTable.km)
      ARI
0.002285344
```

The results are showing the information about cluster means, clustering vector, sum of square by cluster and available components. The fpc package is used to draw the discriminant projection plot using Plotcluster function (Fig3).

The result of plotting the class returned by function application is shown below.

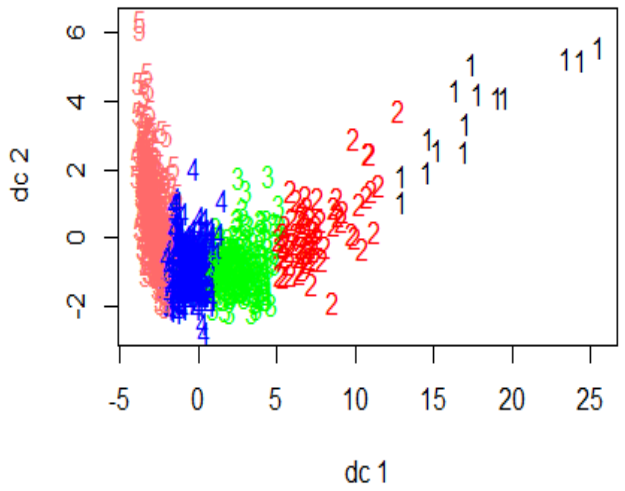

Fig. 3 : Preprocessed K-means plot k=5

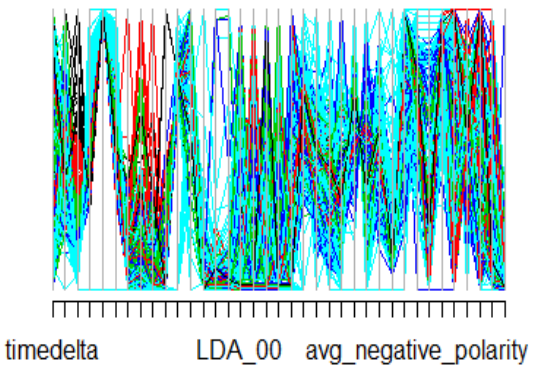

Fig 4 : Parallel coordinators plot

The Figure 4 shows the parallel coordinators plot to see the variables contributed in each cluster.

c) Fuzzy C-Means

Fuzzy c means clustering (FCM), each data point has a fraction of membership to each cluster. This algorithm works iteratively until no further clustering is possible. The membership fraction that minimizes the expected distance to each centroid has to be calculated.

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^{m}\left\|x_i - c_j\right\|^2$$

The algorithm is very similar to K-Means, except that a matrix (row is each data point, column is each centroid, and each cell is the degree of membership) is used.

1. Initialize the membership matrix U
2. Repeat step (3), (4) until converge
3. Compute location of each centroid based on the weighted fraction of its member data point's location.

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} \cdot x_i}{\sum_{i=1}^{N} u_{ij}^{m}}$$

4. Update each cell as follows

$$u_{ij} = \frac{1}{\sum_{k=1}^{C}\left(\frac{\left\|x_i - c_j\right\|}{\left\|x_i - c_k\right\|}\right)^{\frac{2}{m-1}}}$$

Notice that the parameter m is the degree of fuzziness. The output is the matrix with each data point assigned a degree of membership to each centroids.

d) Fuzzy C-Means Implementation Using R

The data repositories used in this paper are The Iris Repository [27]. They are obtained from (http://kdd.ics.uci.edu/). The data set is the fragments of iris flower which is clustered based on the degree assigned by a membership.

The following is the code

```
> library(e1071)
> result <- cmeans(iris[,-5], 3, 100, m=2, method="cmeans")
> plot(iris[,1], iris[,2], col=result$cluster)
> points(result$centers[,c(1,2)], col=1:3, pch=8, cex=2)
```

```
> result$membership[1:3,]
          1           2           3
[1,] 0.001072018 0.002304389 0.9966236
[2,] 0.007498458 0.016651044 0.9758505
[3,] 0.006414909 0.013760502 0.9798246
> table(iris$Species, result$cluster)
            1  2  3
setosa      0  0 50
versicolor  3 47  0
virginica  37 13  0
```

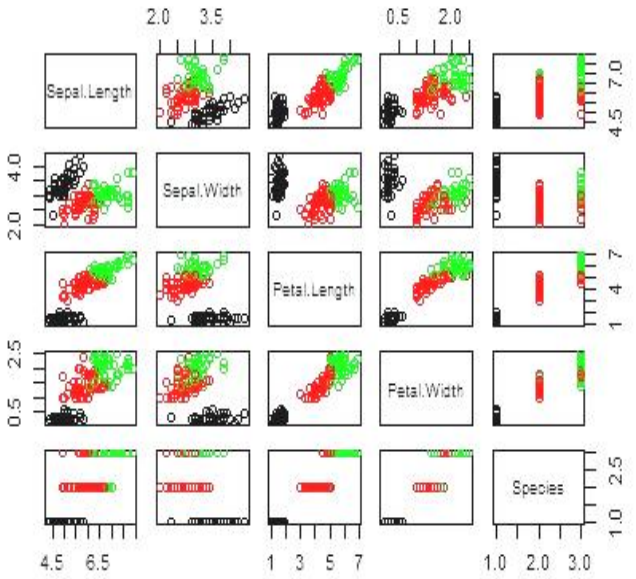

Fig. 5 : Fuzzy C means clustering plot

The observation of the above results and of the Online News Popularity datasets are FCM is mainly dependent on the initial clustering and the computation time is very high for the large data sets. The weight and accuracy are inversely proportional. It is Sensible to noise and membership degree for outliers or noisy points is very low.

e) Hierarchical Clustering

Hierarchical Clustering a series of partitions takes place, which may run from a single cluster containing all objects to n clusters each containing a single object. This clustering builds a cluster hierarchy or, in other words, a tree of clusters, also known as a dendrogram [7]. Hierarchical algorithms can be further categorized into two kinds [3]

i. *Agglomerative clustering*: This clustering starts with *n* clusters and iteratively merges the number of clusters which are most similar objects or clusters, respectively, until only one cluster is remaining ($n \rightarrow 1$).This requires the defining of closest proximity.

ii. *Divisive clustering:* This clustering starts with one cluster and iteratively splits a cluster until singleton clusters of individual points remain, so that the heterogeneity is reduced as far as possible ($1 \rightarrow n$).This requires the decision of splitting at each step.

The Hierarchical Clustering algorithm [6] below takes an n×n distance matrix d input and increasingly gives n different partitions of the data as the tree it outputs result. The largest partition has n single-element clusters, with every element forming its own cluster. The second-largest partition aggregates the two closest clusters from the largest partition, and thus has n-1 clusters. In general, the ith partition combines the two closest clusters from the $(i-1)$th partition and has $(n-i+1)$ clusters. Because of the additional complexity of keeping data in a sorted list or heap, so the time required is $O(m^2 \log m)$ and Space required is $O(m^2)$.
In this approach, it compares all pairs of data points and merges the one with the closest distance.

Algorithm

1: Compute the proximity matrix if necessary
2: repeat
3: Merge the closest two clusters.
4: Update the proximity matrix to reflect the proximity between the new cluster and initial cluster
5: Until only one cluster remains.

The Proximity ($C_i C_j$) of clusters C_i and C_j which are of the size m_i and m_j,respectively is expressed as

$$Proximity\ (c_i, c_j) = \frac{\sum_{x \in c_i} \sum_{y \in c_j} Proximity\ (x,y)}{m_i * m_j}$$

The data set considered is micro RNA expressions. It is actually collected from Fresh paired tumor and control samples from the PAC (Periampullary Carcinoma) patients undergoing Whipple's pan creaticoduodenectomys Data Mining in Health Informatics[15] is an emerging discipline, concerned with developing methods for exploring the unique type of data that come from Health Care database management system. We have also considered the Iris dataset. The code for the implementation is given as follows [16] Figure 6 shows the results of clustering.

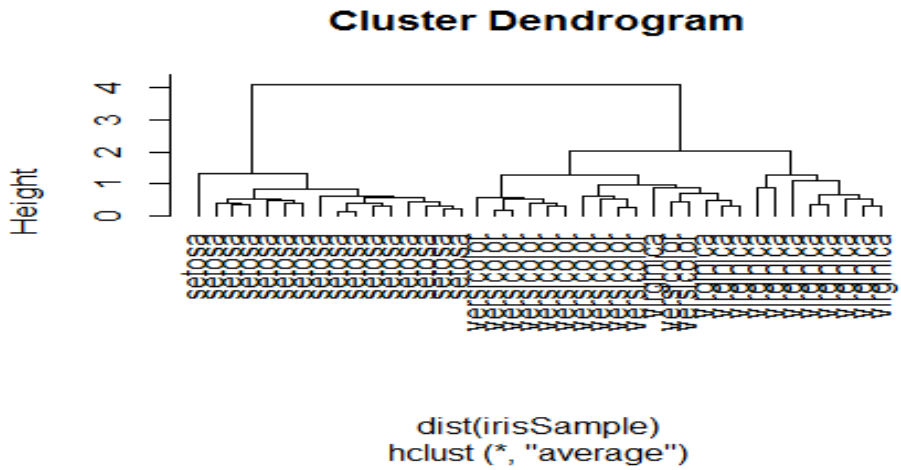

Fig. 6 : Hierarchical Clustering

The Authors [14] have also performed hierarchical clustering on PAC tumors dataset which are distinct from counterpart normal pancreas, normal duodenum, and normal distal CBD and normal ampulla. Unsupervised hierarchical cluster analysis of miRNA Expression profiles were clustered of PAC tumors into pancreatobiliary (n ¼ 23) and intestinal subtypes (n ¼ 17), while normal pancreas (n ¼ 22), normal duodenum (n ¼ 6), normal distal CBD (n ¼ 6) normal ampulla (n ¼ 6) are clustered as different entities (Figure 7).

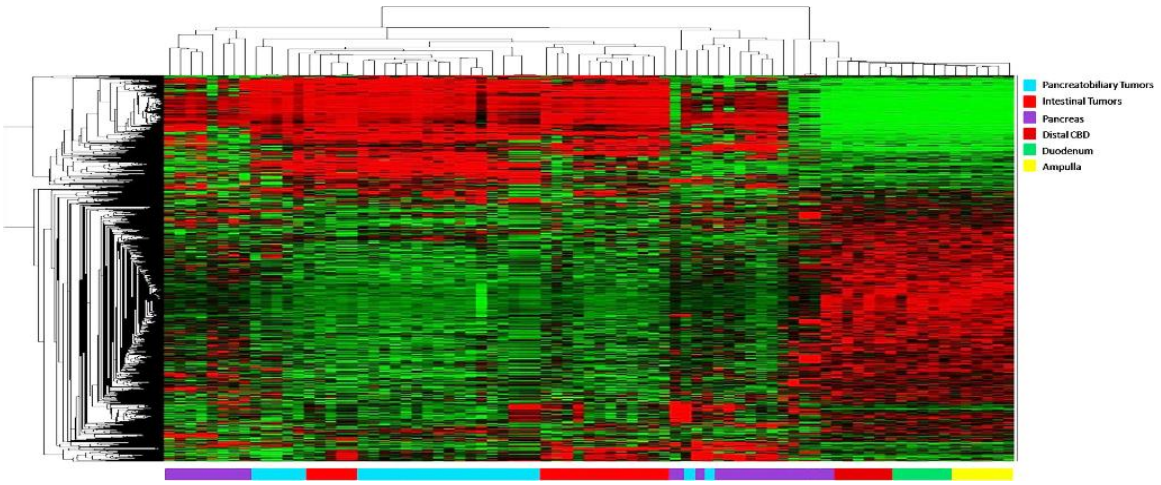

Fig. 7 : The microRNA expression profiles of PAC

IV. Conclusion

We have perceived a comprehensive scan of the k-means, Fuzzy C means and Hierarchical clustering methods using medical research datasets and iris dataset. Using R clustering, Statistical Computing and graphics are represented. All the clustering techniques show ambiguity in clustering noisy data and outliers. The Hierarchical clustering shows good results for small data sets and Fuzzy C means for the voluminous amount of data. K means technique has faster performance but finding the appropriate k value is a big challenge especially in medical research data sets. In continuation to this work we would like to improve clustering analysis by considering the time and accuracy for large data sets using R tool statistics.

V. Acknowledgements

The authors would like to thank Dr. K.S.S Uma Mahesh (M.D Biochemistry) and his lab for sharing their data of PAC tumor samples (microRNA).

References Références Referencias

1. Han J. and Kamber M.: "Data Mining: Concepts and Techniques," Morgan Kaufmann Publishers, San Francisco, 2000.
2. Pang-Ning Tan, Michael Steinbach and Vipin Kumar. Introduction to Data Mining. Addison Wesley; US edition. May 12, 2005. (HC)
3. Data Mining with R: learning by case studies Luis Torgo(kmeans)

4. K. Fernandes, P. Vinagre and P. Cortez. A Proactive Intelligent Decision Support System for Predicting the Popularity of Online News. Proceedings of the 17th EPIA 2015 - Portuguese Conference on Artificial Intelligence, September, Coimbra, Portugal.
5. Jiawei Han and Micheline Kamber, "Data Mining: Concepts and Techniques", Morgan Kaufmann Publishers, second Edition, (2006).
6. Doug Fisher, Optimization and Simplification of Hierarchical Clustering, KDD(HC)
7. Manish Verma, Mauly Srivastava, Neha Chack, Atul Kumar Diswar and Nidhi Gupta, "A Comparative Study of Various Clustering Algorithms in Data Mining", International Journal of Engineering Research and Applications (IJERA) Vol. 2, Issue 3, May-Jun 2012, pp.1379-1384
8. Frank, A. & Asuncion, A. UCI Machine Learning Repository (http:/ / archive. ics. uci. edu/ ml). Irvine, CA: University of California, School of Information and Computer Science 2010.
9. B¨OHM, C., KAILING, K., KRIEGEL, H.-P., AND KR¨OGER, P. 2004. Density connected clustering with local subspace preferences. In Proceedings of the 4th International Conference on Data Mining (ICDM).
10. Robert Gentleman Rafael A. Irizarry Vincent J. Carey Sandrine Dudoit Wolfgang Huber Editors Bioinformatics and Computational Biology Solutions Using R and Bioconductor, Springer
11. R and Data Mining: Examples and Case Studies 1 Yanchang Zhao
12. https://www.r-project.org/
13. Satish Kumar et al Analysis Clustering Techniques in Biological Data with R, International Journal of Computer Science and Information Technologies (IJCSIT), Vol. 6 (2) , 2015
14. KSS Uma Mahes, Kuruva MM , Mitnala S, Rupjyoti T, Venkat RG, Botlagunta S et al. MicroRNA profiling in periampullary carcinoma. Pancreatology 2014; 14: 36-47
15. Abdul Nazeer, K. A., Sebastian M. P., and. Madhu Kumar S.D., 2011 A Heuristic k-Means Algorithm with Better Accuracy and Efficiency for Clustering Health Informatics Data, Journal of Medical Imaging and Health Informatics(American Scientific Publisher) Vol. 1, 66–71.

Global Journal of Computer Science and Technology: C
Software & Data Engineering
Volume 16 Issue 1 Version 1.0 Year 2016
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

A Neuro Fuzzy Algorithm to Compute Software Effort Estimation

By N. Shivakumar, N. Balaji & K. Ananthakumar

Thiagarajar College of Engineering, India

Abstract- Software Effort Estimation is highly important and considered to be a primary activity in software project management. The accurate estimates are conducted in the development of business case in the earlier stages of project management. This accurate prediction helps the investors and customers to identify the total investment and schedule of the project. The project developers define process to estimate the effort more accurately with the available mythologies using the attributes of the project. The algorithmic estimation models are very simple and reliable but not so accurate. The categorical datasets cannot be estimated using the existing techniques. Also the attributes of effort estimation are measured in linguistic values which may leads to confusion. This paper looks in to the accuracy and reliability of a non-algorithmic approach based on adaptive neuro fuzzy logic in the problem of effort estimation. The performance of the proposed method demonstrates that there is a accurate substantiation of the outcomes with the dataset collected from various projects. The results were compared for its accuracy using MRE and MMRE as the metrics. The research idea in the proposed model for effort estimation is based on project domain and attribute which incorporates the model with more competence in augmenting the crux of neural network to exhibit the advances in software estimation.

Keywords: *ANFIS, effort estimation, MRE, MMRE.*

GJCST-C Classification : *I.5.1 I.2.3*

Strictly as per the compliance and regulations of:

A Neuro Fuzzy Algorithm to Compute Software Effort Estimation

N. Shivakumar [α], N. Balaji [σ] & K. Ananthakumar [ρ]

Abstract- Software Effort Estimation is highly important and considered to be a primary activity in software project management. The accurate estimates are conducted in the development of business case in the earlier stages of project management. This accurate prediction helps the investors and customers to identify the total investment and schedule of the project. The project developers define process to estimate the effort more accurately with the available mythologies using the attributes of the project. The algorithmic estimation models are very simple and reliable but not so accurate. The categorical datasets cannot be estimated using the existing techniques. Also the attributes of effort estimation are measured in linguistic values which may leads to confusion. This paper looks in to the accuracy and reliability of a non-algorithmic approach based on adaptive neuro fuzzy logic in the problem of effort estimation. The performance of the proposed method demonstrates that there is a accurate substantiation of the outcomes with the dataset collected from various projects. The results were compared for its accuracy using MRE and MMRE as the metrics. The research idea in the proposed model for effort estimation is based on project domain and attribute which incorporates the model with more competence in augmenting the crux of neural network to exhibit the advances in software estimation.

Keywords: *ANFIS, effort estimation, MRE, MMRE.*

I. Introduction

Achieving software economics in large-scale software development projectsare very important today. Software effort estimation is the process of determining the accurate effort required to maintain or develop a software. It is always an important practical problem in software engineering which is still unsolved. Effort estimates are done in initial stages of software engineering to calculate the effort in person-months required for the software development. Accurate effort estimation helps in planning design construction and transition phases of development and prioritize the components in business case. Unreliable estimates is the main important reason for project failure, which is expressed in 2007 Comp TIA survey of thousand IT professionals, finding that three of the four most-cited causes of IT project failure are due to poor estimation.

Author α : Assistant Professor, Department of CSE, Thiagarajar College of Engineering, Madurai, India. e-mail: shiva@tce.edu
Author σ : Professor, Department of IT, K.L.N College of Engineering, Madurai, India. e-mail: balajin@klnce.edu
Author ρ : P.G. Student, Department of CSE, Thiagarajar College of Engineering, Madurai, India. e-mail: ananthakumar.k@outlook.com

[Rosencrance 2007]. Noticing the importance of reliable effort estimation, the software project management contributors are now focusing on developing models to generate accurate effort of software during the earlier stages of software development. Effort estimation for software projects are categorized as non-algorithmic and algorithmic models. Algorithmic models applies the mathematical computation method and the non-algorithmic estimation uses fuzzy, neural network and other machine learning techniques. The effectiveness of project management will be compromised if the Project managers are uncertain to adapt genuine estimation methodologies

Boehm proposed a method called COCOMO that utilizes some experimental equation to estimate the effort using inputs like Kilo lines of code (KLOC), number of functions and other effort drivers. Neural network sare introduced in effort estimation process mainly for the training and learning from previous data. The model identifies a positive correlation between the dependent (effort) and independent variables (effort drivers). The half of the available data sets can be given for training and the remaining can be used to derive effort. The other techniques of software effort estimation are bottom-up, top-down, analogy estimation and expert judgments.

II. Related Work

Cuauhtemoc [1] provides justification that Fuzzy logic can be used to predict the effort of the small programs based on lines of code obtained from new and changed (N&C) and reused code from small programs developed by 74 programmers. This was used as the input for the fuzzy model for estimating effort and the accuracy of output was compared with the accuracy of Statistical regression model using the comparison criterion Mean Magnitude Error Relative to the estimate.

Shiyna [2] compares the frameworks designed by using fuzzy logic and Neural Networks based on the accuracy of effort estimation. COCOMO NASA dataset had been used as the input for both the frameworks. These frameworks are validated using the parameters MMRE (Mean Magnitude of Relative Error) and Pred (Prediction Accuracy). The results show that Fuzzy Logic based framework works better when compared to the Neural Network framework.

Ochodek [3] proposed the usage of Use Case Points (UCP) method to estimate the effort based on the use case model and two adjustment factors (With or Without Unadjusted Actor Weights). The cross-validation procedure has been used to compare the variants of adjustment factors. A group of 14 projects is considered as input which are used to arrive at a conclusion that the UCP method can be simplified without the use of adjustment factors.

Iman [4] compares the software effort estimation computed by the conventional methods like function points, regression models and COCOMO with the model designed using fuzzy logic. The parameter Mean Magnitude of Relative Error (MMRE) is used to compute the accuracy of the considered methods.

Anjana Bawa [5] explains the usage of Artificial Neural Networks to estimate the project effort as it is capable of learning from the previous data. The machine learning algorithms, Back-Propagation and Cascade Correlation are used to learn and classify the dataset and hence estimate the effort using the Neural Networks.

By analyzing the previous work, it is evident that fuzzy logic is better than the conventional methods of effort estimation. By using the package points the complexity of estimating the lines of code for the considered software is reduced. By using the factor refinement, the time taken to compute the effort is less compared to the previous method where 15 attributes where obtained from the programmer to compute the effort.

III. Proposed Approach

We have considered 93 instances of NASA historical project data and also investigated and gathered thirty projects from many case studies and experiments [11][12][14][15] with consists of 15 attributes and actual effort along with domain, area of work, Size. The fifteen attributes art converted to three index valued labelled as Human Perception and Performance Index (HPPI), Machine Requirement and Performance Index (MRPI), Process Requirement and Performance Index (PRPI). Adaptive Neuro-Fuzzy model (The Figure 1) for software development effort estimation is perfect in the learning and good interpretability. Artificial neural networks are made up of processing units in a parallel manner called as neurons these neurons are inter linked by connections. The input for this model is six grouped attributes. Each attribute represents one factor which leads to the development effort. Table 1 describes the refinement of attributes in such a way that 15 effort multipliers are grouped in 3 clusters of refined attributes whose values are obtained from the software project developer.

Table 1: Refined Attributes

REFINED ATTRIBUTES	EFFORT MULTIPLIERS
Human Perception and Performance Index (HPPI)	1. Analyst Efficiency 2. Programmer Efficiency 3. Application Maturity 4. Modern Programming Practices 5. Use of Software tools 6. Virtual Machine Experience 7. Language Experience
Machine Requirement and Performance Index (MRPI)	8. Time Constraint for CPU 9. Turnaround Time 10. Machine Volatility
Process Requirement and Performance Index (PRPI)	11. Process Complexity 12. Storage Space Requirement 13. Schedule Constraint 14. Database Size 15. Required Software Reliability

This proposed model consolidates neural networks and fuzzy-logic principles in a combined ANFIS framework. This inference system correlates to the learning capability of fuzzy IF THEN rules to approximate the nonlinear functions.

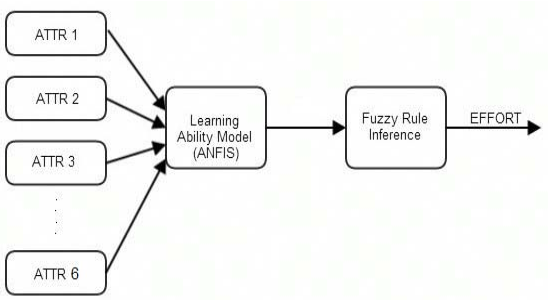

Figure 1 : Adaptive Neuro-Fuzzy model

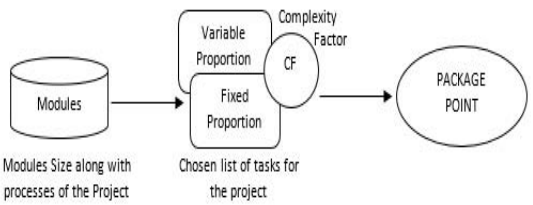

Figure 2 : Package Point Process

a) Package Points

The package point (figure 2) provides an alternate way to estimate the size that needs to be applied in a software project. Unlike function points and class points approach, the package points proves to be highly efficient in aiding to estimate effort and it is proven to work well with the ERP projects. In order to compute the package points, inputs are obtained

namely, scope, tasks and complexities. Package points are defined by standardizing the number of modules in the project primarily. Then the tasks to complete the modules are prioritized and defined by the client as per their requirement. Finally the complexity factors are loaded to arrive at the package point for the considered module.

b) Validation in ANFIS Model

ANFIS is a hybrid supervised method which adopts a hybrid learning algorithm to determine the parameters for fuzzy inference systems. It utilizes both least-squares method and propagation gradient descent method. This is used for training FIS membership function parameters to examine the given training data set. ANFIS can be executed using an optional argument for model validation. This is called as checking model for over fitting. The argument used for this is called as checking data set.

c) Fuzzy Rule Inference

Fuzzy rules are generated using package points, domain, type, Human Perception and Performance Index (HPPI), Machine Requirement and Performance Index (MRPI), Process Requirement and Performance Index (PRPI) and Actual effort. A fuzzy set is illustrated using a membership function that relates with every point in the fuzzy set that comprises of numbers in the interval [0, 1], known as degree or grade of membership. Membership function used in this research work is Triangular Membership Function. Triangular Fuzzy Number (TFN) is defined using a triplet (α, m, β), where m denotes modal value, α and βsignify the right and left boundary correspondingly and is expressed as:

$$\mu(x) = \begin{cases} 0, & x \le \alpha \\ \frac{x-\alpha}{m-\alpha}, & \alpha \le x \le m \\ \frac{\beta-x}{\beta-m}, & m \le x \le \beta \\ 0, & x \ge \beta \end{cases} \quad (1)$$

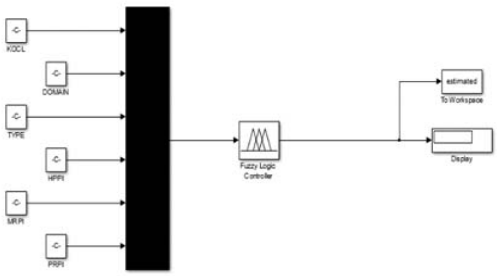

Figure 3 : Fuzzy Rules generated

The equivalent characteristics of the rules the most significant aspect of fuzzy logic systems. Instead of sharp swapping among modes based on breakpoints, logic flows efficiently from sections where the system's performance is governed by one rule or another.

Defuzzification converts from fuzzy to crisp conversions. The process converts the fuzzy value to the estimated value for the single data set. This is similar to a "rounding off" method. Defuzzification converts the collection of membership function data to a single sealer quantity in corresponding membership degrees. It is characterized instructure of rules that convert variables to a fuzzy result, that is, the outcome is defined in terms of membership in fuzzy sets.

$$R_{alv} = \begin{cases} \mu^* w_1 & 0 < \text{ALV} \le 1 \\ \mu^* w_1 + (1-\mu) * w_2 & 1 < \text{ALV} \le 2 \\ \mu^* w_1 + (1-\mu) * w_1 & 2 < \text{ALV} \le 3.5 \\ \mu^* w_2 + (1-\mu)\, w_3 & 3.5 < \text{ALV} \le 5 \\ \mu^* w_3 + (1-\mu)\, w_2 & 5 < \text{ALV} \le 6.5 \\ \mu^* w_3 + (1-\mu)\, w_4 & 6\,5 < \text{ALV} \le 8 \\ w_4 & \text{ALV} > 8 \end{cases} \quad (2)$$

IV. Experimental Design

a) Evaluation Criteria

1. Mean Magnitude Relative Error (MRE), is an error ratio between the absolute deviations of prediction to the actual effort in each of the referred project

$$MRE = |(Actual_i - Estimated_i)|/(Actual_i) \quad (3)$$

2. Mean Magnitude Relative Error (MMRE) is the average value of MER of all the referred projects

$$MMRE = 1/n \sum MREi \quad (4)$$

The Table 2 indicates the Package point and its subsequent domain with its actual effort and the non-algorithmic estimated effort using the proposed method. The effort in the dataset is compared with this estimated effort

Table 2 : Domain based Metrics

Area Of Domain	Package Point	Actual Effort	Estimated Effort
Avionics monitoring	25.9	117.6	138.4
Mission planning	31.5	60	112
Simulation	66.6	352.8	402
Monitor Control	70	458	561
Real Data Processing	177.9	124	397
Communications	240	192	322
Batch Data Processing	25.9	117.6	119
Data Capture	31.5	60	67.7
Launch Processing	66.6	352.8	360.9
Application Ground	70	458	459
Utility	177.9	124	128
Operating Systems	240	192	193.2

The preceding list of algorithmic models are tested including KLOC (Kilo Lines Of Code) value from the data sets and efforts were estimated and these estimates are compared with the results obtained from Adaptive Neuro Fuzzy model.

- Halstead Model- This model developed by Halstead, concerning the supplied lines of source code from the programmer and formulates a relation,

$$Effort = 0.7 * (KLOC)^{2} \quad (5)$$

- Bailey-Basili Model - This model developed by Bailey-Basili, between delivered lines of source code and formulates a relation,

$$Effort = 5.5 * (KLOC)^{1.16} \quad (6)$$

- Doty Model - This model developed by Doty, between delivered lines of source code and formulates a relation,

$$Effort = 5.288 * (KLOC)^{1.047} \quad (7)$$

- COCOMO - It was the first model suggested by Barry Boehm. This model has been widely accepted in practice. In the COCOMO model, the code-size S is given in thousand LOC (KLOC) and Effort is in person-month.

$$Effort = a * (KLOC)^{b} \quad (8)$$

Where a, b are complexity factors. This model uses three sets of a, b depending on the complexity of the software. The basic COCOMO model is simple and easy to use.

- COCOMO II - It comprises of three variants, namely, early design model, Application composition model, and Post architecture model. This is an augmentation of intermediate COCOMO model and defined as,

$$Effort = 2.9 * (KLOC)^{1.10} \quad (9)$$

b) Numerical Results

This section reports the experimental results of predictions obtained from soft computing models and other algorithmic models. The actual effort existing in the dataset is associated with the estimated effort and finally Mean Relative Error is calculated. The following graphs (figure 4) shows the Mean Relative Error variations in Doty, Bailey, COCOMO I and COCOMO II, Halstead and Neuro Fuzzy models.

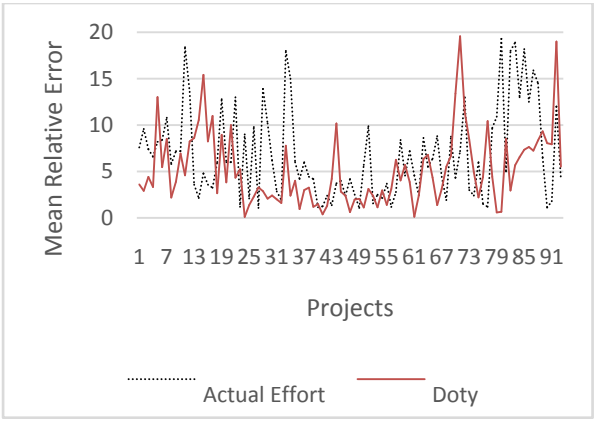

Figure 4 : Doty Estimation Vs Actual Effort

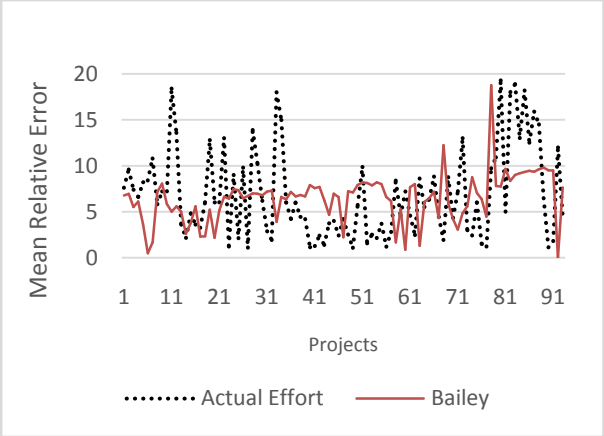

Figure 5 : Bailey Estimation Vs Actual Effort

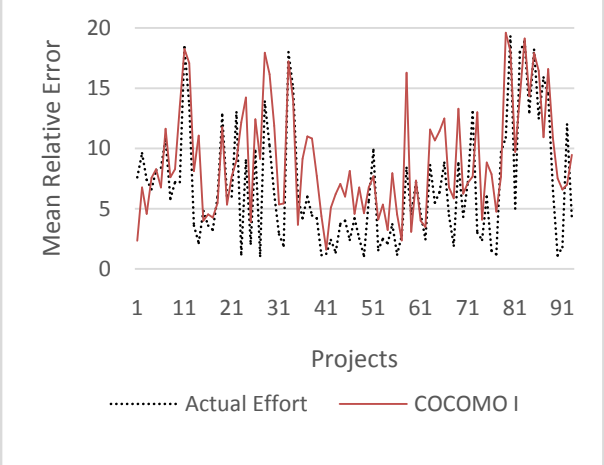

Figure 6 : COCOMO I Estimation Vs Actual Effort

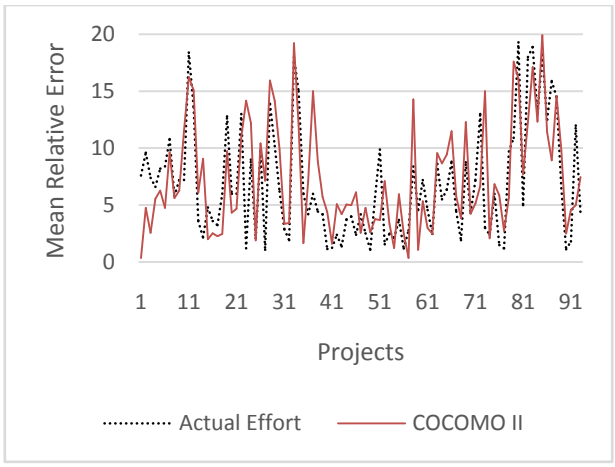

Figure 7 : COCOMO II Estimation Vs Actual Effort

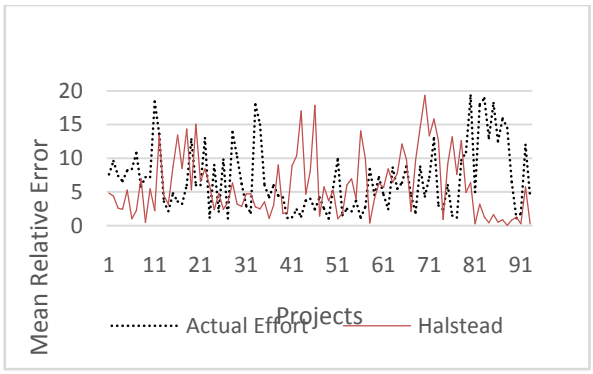

Figure 8 : Halstead Estimation Vs Actual Effort

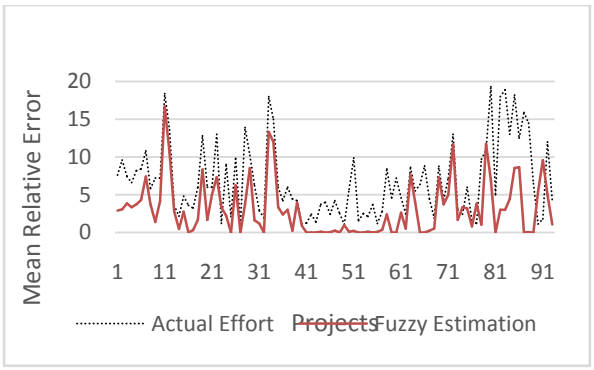

Figure 9 : Fuzzy Based Estimation Vs Actual Effort

The following table summarizes the MMRE value of all the algorithmic and non-algorithmic models discussed above for all the referred projects.

Table 3 : MMRE comparison

	Doty	Bailey	Halstead	COCOMO I	COCOMO II	Fuzzy Estimation
MMRE	5.15	6.47	6.06	9.14	7.54	3.11

V. Conclusion

Early effort estimation in software development lifecycle is an important activity for project planning and resource allocation. This research work proposes an efficient model in estimating the software effort. The outcomes of the estimation obtained using the direct algorithmic methods indicates the divergence between the actual and the estimated effort. The outcome of non-algorithmic method comprising of the adaptive neuro technique based estimation decreases the Mean Magnitude of Relative Error (MMRE). Hence the examination of effort from algorithmic method and non-algorithmic method prove that adaptive neuro fuzzy based estimation is more efficient than the algorithmic methods for the estimation process. The success of estimation depends upon the accuracy and stability of the method in various measures. Future work is planned to investigate the clustering algorithms in estimation process and apply Neuro Fuzzy model on large datasets.

References Références Referencias

1. Cuauhtemoc Lopez-Martin, "A fuzzy logic model for predicting the development effort of short scale programs based upon two independent variables", Applied Soft Computing, Vol.11, 2010.
2. Shiyna Kumar, Vinay Chopra, "Neural Network and Fuzzy Logic based framework for Software Development Effort Estimation", International Journal of Advanced Research in Computer Science and Software Engineering, Vol.3, No. 5, 2013.
3. M. Ochodek, J.Nawrocki and K.Kwarciak, "Simplifying effort estimation based on Use Case Points", Information and Software Technology, Vol.53, 2011.
4. Iman Attarzadeh and Siew Hock Ow, "Software Development Effort Estimation Based on a New Fuzzy Logic Model", International Journal of Computer Theory and Engineering, Vol. 1, No. 4, 2009.
5. Anjana Bawa and Rama Chawla, "Experimental Analysis of Effort Estimation Using Artificial Neural

Network", International Journal of Electronics and Computer Science Engineering, Vol.1, No.3, 2011.
6. Mohammad Azzeh, Daniel Neagu and Peter I. Cowling, "Analogy-based software effort estimation using Fuzzy numbers", The Journal of Systems and Software, Vol.84, 2011.
7. Sun-Jen Huang, Nan-Hsing Chiu and Li-Wei Chen, "Integration of the grey relational analysis with genetic algorithm for software effort estimation", European Journal of Operational Research, Vol.188, 2007.
8. M. Kazemifard, A. Zaeri, N. Ghasem-Aghaee, M.A. Nematbakhsh and F. Mardukhi , "Fuzzy Emotional COCOMO II Software Cost Estimation (FECSCE) using Multi-Agent Systems", Applied Soft Computing, Vol.11, 2011.
9. Emad A. El-Sebakhy, "Functional networks as a novel data mining paradigm in forecasting software development efforts", Expert Systems with Applications, Vol.38, 2011.
10. Magne Jørgensen, "Contrasting ideal and realistic conditions as a means to improve judgment-based software development effort estimation", Information and Software Technology, Vol.53, 2011.
11. Bente Anda, Endre Angelvik, and Kirsten Ribu,"Improving Estimation Practices by Applying Use Case Models", product focused software process improvement, Springer,2002. Kirsten Ribu, "Estimating Object-Oriented Software Projects with Use Cases, University of Oslo. 2001
12. AP Subriadi, Sholiq, A P Ningrum. "Critical review of the effort rate value in use case point method for estimating software development effort" Journal of Theoretical and Applied Information Technology, January 2014
13. M. Azzeha, Ali Bou Nassifb, Leandro L. Minkuc, "An empirical evaluation of ensemble adjustment methods for analogy-based effort estimation" Journal of Systems and Software, may 2015
14. A.B Nassif, L.F. Capretz, "Software Effort Estimation in the Early Stages of the Software Life Cycle Using a Cascade Correlation Neural Network Model" 13th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel & Distributed Computing (SNPD), 2012
15. Nassif, A.B.; Capretz, L.F.; Ho, D. "Estimating Software Effort Using an ANN Model Based on Use Case Points", Machine Learning and Applications (ICMLA), 2012 11th International Conference on, On page(s): 42 - 47 Volume: 2, 12-15 Dec. 2012
16. El Bajta, M. "Analogy-Based Software Development Effort Estimation in Global Software Development", Global Software Engineering Workshops (ICGSEW), 2015 IEEE 10th International Conference on, on page(s): 51 – 54.

Global Journals Inc. (US) Guidelines Handbook 2016

www.GlobalJournals.org

FELLOW OF ASSOCIATION OF RESEARCH SOCIETY IN COMPUTING (FARSC)

Global Journals Incorporate (USA) is accredited by Open Association of Research Society (OARS), U.S.A and in turn, awards "FARSC" title to individuals. The 'FARSC' title is accorded to a selected professional after the approval of the Editor-in-Chief/Editorial Board Members/Dean.

- The "FARSC" is a dignified title which is accorded to a person's name viz. Dr. John E. Hall, Ph.D., FARSC or William Walldroff, M.S., FARSC.

FARSC accrediting is an honor. It authenticates your research activities. After recognition as FARSC, you can add 'FARSC' title with your name as you use this recognition as additional suffix to your status. This will definitely enhance and add more value and repute to your name. You may use it on your professional Counseling Materials such as CV, Resume, and Visiting Card etc.

The following benefits can be availed by you only for next three years from the date of certification:

FARSC designated members are entitled to avail a 40% discount while publishing their research papers (of a single author) with Global Journals Incorporation (USA), if the same is accepted by Editorial Board/Peer Reviewers. If you are a main author or co-author in case of multiple authors, you will be entitled to avail discount of 10%.

Once FARSC title is accorded, the Fellow is authorized to organize a symposium/seminar/conference on behalf of Global Journal Incorporation (USA).The Fellow can also participate in conference/seminar/symposium organized by another institution as representative of Global Journal. In both the cases, it is mandatory for him to discuss with us and obtain our consent.

You may join as member of the Editorial Board of Global Journals Incorporation (USA) after successful completion of three years as Fellow and as Peer Reviewer. In addition, it is also desirable that you should organize seminar/symposium/conference at least once.

We shall provide you intimation regarding launching of e-version of journal of your stream time to time. This may be utilized in your library for the enrichment of knowledge of your students as well as it can also be helpful for the concerned faculty members.

The FARSC can go through standards of OARS. You can also play vital role if you have any suggestions so that proper amendment can take place to improve the same for the benefit of entire research community.

As FARSC, you will be given a renowned, secure and free professional email address with 100 GB of space e.g. johnhall@globaljournals.org. This will include Webmail, Spam Assassin, Email Forwarders, Auto-Responders, Email Delivery Route tracing, etc.

The FARSC will be eligible for a free application of standardization of their researches. Standardization of research will be subject to acceptability within stipulated norms as the next step after publishing in a journal. We shall depute a team of specialized research professionals who will render their services for elevating your researches to next higher level, which is worldwide open standardization.

The FARSC member can apply for grading and certification of standards of their educational and Institutional Degrees to Open Association of Research, Society U.S.A. Once you are designated as FARSC, you may send us a scanned copy of all of your credentials. OARS will verify, grade and certify them. This will be based on your academic records, quality of research papers published by you, and some more criteria. After certification of all your credentials by OARS, they will be published on your Fellow Profile link on website https://associationofresearch.org which will be helpful to upgrade the dignity.

The FARSC members can avail the benefits of free research podcasting in Global Research Radio with their research documents. After publishing the work, (including published elsewhere worldwide with proper authorization) you can upload your research paper with your recorded voice or you can utilize chargeable services of our professional RJs to record your paper in their voice on request.

The FARSC member also entitled to get the benefits of free research podcasting of their research documents through video clips. We can also streamline your conference videos and display your slides/ online slides and online research video clips at reasonable charges, on request.

The FARSC is eligible to earn from sales proceeds of his/her researches/reference/review Books or literature, while publishing with Global Journals. The FARSC can decide whether he/she would like to publish his/her research in a closed manner. In this case, whenever readers purchase that individual research paper for reading, maximum 60% of its profit earned as royalty by Global Journals, will be credited to his/her bank account. The entire entitled amount will be credited to his/her bank account exceeding limit of minimum fixed balance. There is no minimum time limit for collection. The FARSC member can decide its price and we can help in making the right decision.

The FARSC member is eligible to join as a paid peer reviewer at Global Journals Incorporation (USA) and can get remuneration of 15% of author fees, taken from the author of a respective paper. After reviewing 5 or more papers you can request to transfer the amount to your bank account.

MEMBER OF ASSOCIATION OF RESEARCH SOCIETY IN COMPUTING (MARSC)

The ' MARSC ' title is accorded to a selected professional after the approval of the Editor-in-Chief / Editorial Board Members/Dean.
The “MARSC” is a dignified ornament which is accorded to a person’s name viz. Dr. John E. Hall, Ph.D., MARSC or William Walldroff, M.S., MARSC.

MARSC accrediting is an honor. It authenticates your research activities. After becoming MARSC, you can add 'MARSC' title with your name as you use this recognition as additional suffix to your status. This will definitely enhance and add more value and repute to your name. You may use it on your professional Counseling Materials such as CV, Resume, Visiting Card and Name Plate etc.

The following benefitscan be availed by you only for next three years from the date of certification.

MARSC designated members are entitled to avail a 25% discount while publishing their research papers (of a single author) in Global Journals Inc., if the same is accepted by our Editorial Board and Peer Reviewers. If you are a main author or co-author of a group of authors, you will get discount of 10%.

As MARSC, you will be given a renowned, secure and free professional email address with 30 GB of space e.g. johnhall@globaljournals.org. This will include Webmail, Spam Assassin, Email Forwarders, Auto-Responders, Email Delivery Route tracing, etc.

We shall provide you intimation regarding launching of e-version of journal of your stream time to time. This may be utilized in your library for the enrichment of knowledge of your students as well as it can also be helpful for the concerned faculty members.

The MARSC member can apply for approval, grading and certification of standards of their educational and Institutional Degrees to Open Association of Research, Society U.S.A.

Once you are designated as MARSC, you may send us a scanned copy of all of your credentials. OARS will verify, grade and certify them. This will be based on your academic records, quality of research papers published by you, and some more criteria.

It is mandatory to read all terms and conditions carefully.

AUXILIARY MEMBERSHIPS

Institutional Fellow of Open Association of Research Society (USA)-OARS (USA)

Global Journals Incorporation (USA) is accredited by Open Association of Research Society, U.S.A (OARS) and in turn, affiliates research institutions as "Institutional Fellow of Open Association of Research Society" (IFOARS).

The "FARSC" is a dignified title which is accorded to a person's name viz. Dr. John E. Hall, Ph.D., FARSC or William Walldroff, M.S., FARSC.

The IFOARS institution is entitled to form a Board comprised of one Chairperson and three to five board members preferably from different streams. The Board will be recognized as "Institutional Board of Open Association of Research Society"-(IBOARS).

The Institute will be entitled to following benefits:

The IBOARS can initially review research papers of their institute and recommend them to publish with respective journal of Global Journals. It can also review the papers of other institutions after obtaining our consent. The second review will be done by peer reviewer of Global Journals Incorporation (USA) The Board is at liberty to appoint a peer reviewer with the approval of chairperson after consulting us.

The author fees of such paper may be waived off up to 40%.

The Global Journals Incorporation (USA) at its discretion can also refer double blind peer reviewed paper at their end to the board for the verification and to get recommendation for final stage of acceptance of publication.

The IBOARS can organize symposium/seminar/conference in their country on behalf of Global Journals Incorporation (USA)-OARS (USA). The terms and conditions can be discussed separately.

The Board can also play vital role by exploring and giving valuable suggestions regarding the Standards of "Open Association of Research Society, U.S.A (OARS)" so that proper amendment can take place for the benefit of entire research community. We shall provide details of particular standard only on receipt of request from the Board.

The board members can also join us as Individual Fellow with 40% discount on total fees applicable to Individual Fellow. They will be entitled to avail all the benefits as declared. Please visit Individual Fellow-sub menu of GlobalJournals.org to have more relevant details.

We shall provide you intimation regarding launching of e-version of journal of your stream time to time. This may be utilized in your library for the enrichment of knowledge of your students as well as it can also be helpful for the concerned faculty members.

After nomination of your institution as "Institutional Fellow" and constantly functioning successfully for one year, we can consider giving recognition to your institute to function as Regional/Zonal office on our behalf.
The board can also take up the additional allied activities for betterment after our consultation.

The following entitlements are applicable to individual Fellows:

Open Association of Research Society, U.S.A (OARS) By-laws states that an individual Fellow may use the designations as applicable, or the corresponding initials. The Credentials of individual Fellow and Associate designations signify that the individual has gained knowledge of the fundamental concepts. One is magnanimous and proficient in an expertise course covering the professional code of conduct, and follows recognized standards of practice.

Open Association of Research Society (US)/ Global Journals Incorporation (USA), as described in Corporate Statements, are educational, research publishing and professional membership organizations. Achieving our individual Fellow or Associate status is based mainly on meeting stated educational research requirements.

Disbursement of 40% Royalty earned through Global Journals : Researcher = 50%, Peer Reviewer = 37.50%, Institution = 12.50% E.g. Out of 40%, the 20% benefit should be passed on to researcher, 15 % benefit towards remuneration should be given to a reviewer and remaining 5% is to be retained by the institution.

We shall provide print version of 12 issues of any three journals [as per your requirement] out of our 38 journals worth $ 2376 USD.

Other:

The individual Fellow and Associate designations accredited by Open Association of Research Society (US) credentials signify guarantees following achievements:

- The professional accredited with Fellow honor, is entitled to various benefits viz. name, fame, honor, regular flow of income, secured bright future, social status etc.

- In addition to above, if one is single author, then entitled to 40% discount on publishing research paper and can get 10%discount if one is co-author or main author among group of authors.
- The Fellow can organize symposium/seminar/conference on behalf of Global Journals Incorporation (USA) and he/she can also attend the same organized by other institutes on behalf of Global Journals.
- The Fellow can become member of Editorial Board Member after completing 3yrs.
- The Fellow can earn 60% of sales proceeds from the sale of reference/review books/literature/publishing of research paper.
- Fellow can also join as paid peer reviewer and earn 15% remuneration of author charges and can also get an opportunity to join as member of the Editorial Board of Global Journals Incorporation (USA)
- • This individual has learned the basic methods of applying those concepts and techniques to common challenging situations. This individual has further demonstrated an in–depth understanding of the application of suitable techniques to a particular area of research practice.

Note :

"

- In future, if the board feels the necessity to change any board member, the same can be done with the consent of the chairperson along with anyone board member without our approval.
- In case, the chairperson needs to be replaced then consent of 2/3rd board members are required and they are also required to jointly pass the resolution copy of which should be sent to us. In such case, it will be compulsory to obtain our approval before replacement.
- In case of "Difference of Opinion [if any]" among the Board members, our decision will be final and binding to everyone.

"

Process of submission of Research Paper

The Area or field of specialization may or may not be of any category as mentioned in 'Scope of Journal' menu of the GlobalJournals.org website. There are 37 Research Journal categorized with Six parental Journals GJCST, GJMR, GJRE, GJMBR, GJSFR, GJHSS. For Authors should prefer the mentioned categories. There are three widely used systems UDC, DDC and LCC. The details are available as 'Knowledge Abstract' at Home page. The major advantage of this coding is that, the research work will be exposed to and shared with all over the world as we are being abstracted and indexed worldwide.

The paper should be in proper format. The format can be downloaded from first page of 'Author Guideline' Menu. The Author is expected to follow the general rules as mentioned in this menu. The paper should be written in MS-Word Format (*.DOC,*.DOCX).

The Author can submit the paper either online or offline. The authors should prefer online submission.Online Submission: There are three ways to submit your paper:

(A) (I) First, register yourself using top right corner of Home page then Login. If you are already registered, then login using your username and password.

(II) Choose corresponding Journal.

(III) Click 'Submit Manuscript'. Fill required information and Upload the paper.

(B) If you are using Internet Explorer, then Direct Submission through Homepage is also available.

(C) If these two are not convenient, and then email the paper directly to dean@globaljournals.org.

Offline Submission: Author can send the typed form of paper by Post. However, online submission should be preferred.

Preferred Author Guidelines

MANUSCRIPT STYLE INSTRUCTION (Must be strictly followed)

Page Size: 8.27" X 11'"

- Left Margin: 0.65
- Right Margin: 0.65
- Top Margin: 0.75
- Bottom Margin: 0.75
- Font type of all text should be Swis 721 Lt BT.
- Paper Title should be of Font Size 24 with one Column section.
- Author Name in Font Size of 11 with one column as of Title.
- Abstract Font size of 9 Bold, "Abstract" word in Italic Bold.
- Main Text: Font size 10 with justified two columns section
- Two Column with Equal Column with of 3.38 and Gaping of .2
- First Character must be three lines Drop capped.
- Paragraph before Spacing of 1 pt and After of 0 pt.
- Line Spacing of 1 pt
- Large Images must be in One Column
- Numbering of First Main Headings (Heading 1) must be in Roman Letters, Capital Letter, and Font Size of 10.
- Numbering of Second Main Headings (Heading 2) must be in Alphabets, Italic, and Font Size of 10.

You can use your own standard format also.
Author Guidelines:

1. General,

2. Ethical Guidelines,

3. Submission of Manuscripts,

4. Manuscript's Category,

5. Structure and Format of Manuscript,

6. After Acceptance.

1. GENERAL

Before submitting your research paper, one is advised to go through the details as mentioned in following heads. It will be beneficial, while peer reviewer justify your paper for publication.

Scope

The Global Journals Inc. (US) welcome the submission of original paper, review paper, survey article relevant to the all the streams of Philosophy and knowledge. The Global Journals Inc. (US) is parental platform for Global Journal of Computer Science and Technology, Researches in Engineering, Medical Research, Science Frontier Research, Human Social Science, Management, and Business organization. The choice of specific field can be done otherwise as following in Abstracting and Indexing Page on this Website. As the all Global

Journals Inc. (US) are being abstracted and indexed (in process) by most of the reputed organizations. Topics of only narrow interest will not be accepted unless they have wider potential or consequences.

2. ETHICAL GUIDELINES

Authors should follow the ethical guidelines as mentioned below for publication of research paper and research activities.

Papers are accepted on strict understanding that the material in whole or in part has not been, nor is being, considered for publication elsewhere. If the paper once accepted by Global Journals Inc. (US) and Editorial Board, will become the copyright of the Global Journals Inc. (US).

Authorship: The authors and coauthors should have active contribution to conception design, analysis and interpretation of findings. They should critically review the contents and drafting of the paper. All should approve the final version of the paper before submission

The Global Journals Inc. (US) follows the definition of authorship set up by the Global Academy of Research and Development. According to the Global Academy of R&D authorship, criteria must be based on:

1) Substantial contributions to conception and acquisition of data, analysis and interpretation of the findings.

2) Drafting the paper and revising it critically regarding important academic content.

3) Final approval of the version of the paper to be published.

All authors should have been credited according to their appropriate contribution in research activity and preparing paper. Contributors who do not match the criteria as authors may be mentioned under Acknowledgement.

Acknowledgements: Contributors to the research other than authors credited should be mentioned under acknowledgement. The specifications of the source of funding for the research if appropriate can be included. Suppliers of resources may be mentioned along with address.

Appeal of Decision: The Editorial Board's decision on publication of the paper is final and cannot be appealed elsewhere.

Permissions: It is the author's responsibility to have prior permission if all or parts of earlier published illustrations are used in this paper.

Please mention proper reference and appropriate acknowledgements wherever expected.

If all or parts of previously published illustrations are used, permission must be taken from the copyright holder concerned. It is the author's responsibility to take these in writing.

Approval for reproduction/modification of any information (including figures and tables) published elsewhere must be obtained by the authors/copyright holders before submission of the manuscript. Contributors (Authors) are responsible for any copyright fee involved.

3. SUBMISSION OF MANUSCRIPTS

Manuscripts should be uploaded via this online submission page. The online submission is most efficient method for submission of papers, as it enables rapid distribution of manuscripts and consequently speeds up the review procedure. It also enables authors to know the status of their own manuscripts by emailing us. Complete instructions for submitting a paper is available below.

Manuscript submission is a systematic procedure and little preparation is required beyond having all parts of your manuscript in a given format and a computer with an Internet connection and a Web browser. Full help and instructions are provided on-screen. As an author, you will be prompted for login and manuscript details as Field of Paper and then to upload your manuscript file(s) according to the instructions.

To avoid postal delays, all transaction is preferred by e-mail. A finished manuscript submission is confirmed by e-mail immediately and your paper enters the editorial process with no postal delays. When a conclusion is made about the publication of your paper by our Editorial Board, revisions can be submitted online with the same procedure, with an occasion to view and respond to all comments.

Complete support for both authors and co-author is provided.

4. MANUSCRIPT'S CATEGORY

Based on potential and nature, the manuscript can be categorized under the following heads:

Original research paper: Such papers are reports of high-level significant original research work.

Review papers: These are concise, significant but helpful and decisive topics for young researchers.

Research articles: These are handled with small investigation and applications.

Research letters: The letters are small and concise comments on previously published matters.

5. STRUCTURE AND FORMAT OF MANUSCRIPT

The recommended size of original research paper is less than seven thousand words, review papers fewer than seven thousands words also.Preparation of research paper or how to write research paper, are major hurdle, while writing manuscript. The research articles and research letters should be fewer than three thousand words, the structure original research paper; sometime review paper should be as follows:

Papers: These are reports of significant research (typically less than 7000 words equivalent, including tables, figures, references), and comprise:

(a)Title should be relevant and commensurate with the theme of the paper.

(b) A brief Summary, "Abstract" (less than 150 words) containing the major results and conclusions.

(c) Up to ten keywords, that precisely identifies the paper's subject, purpose, and focus.

(d) An Introduction, giving necessary background excluding subheadings; objectives must be clearly declared.

(e) Resources and techniques with sufficient complete experimental details (wherever possible by reference) to permit repetition; sources of information must be given and numerical methods must be specified by reference, unless non-standard.

(f) Results should be presented concisely, by well-designed tables and/or figures; the same data may not be used in both; suitable statistical data should be given. All data must be obtained with attention to numerical detail in the planning stage. As reproduced design has been recognized to be important to experiments for a considerable time, the Editor has decided that any paper that appears not to have adequate numerical treatments of the data will be returned un-refereed;

(g) Discussion should cover the implications and consequences, not just recapitulating the results; conclusions should be summarizing.

(h) Brief Acknowledgements.

(i) References in the proper form.

Authors should very cautiously consider the preparation of papers to ensure that they communicate efficiently. Papers are much more likely to be accepted, if they are cautiously designed and laid out, contain few or no errors, are summarizing, and be conventional to the approach and instructions. They will in addition, be published with much less delays than those that require much technical and editorial correction.

The Editorial Board reserves the right to make literary corrections and to make suggestions to improve briefness.

It is vital, that authors take care in submitting a manuscript that is written in simple language and adheres to published guidelines.

Format

Language: The language of publication is UK English. Authors, for whom English is a second language, must have their manuscript efficiently edited by an English-speaking person before submission to make sure that, the English is of high excellence. It is preferable, that manuscripts should be professionally edited.

Standard Usage, Abbreviations, and Units: Spelling and hyphenation should be conventional to The Concise Oxford English Dictionary. Statistics and measurements should at all times be given in figures, e.g. 16 min, except for when the number begins a sentence. When the number does not refer to a unit of measurement it should be spelt in full unless, it is 160 or greater.

Abbreviations supposed to be used carefully. The abbreviated name or expression is supposed to be cited in full at first usage, followed by the conventional abbreviation in parentheses.

Metric SI units are supposed to generally be used excluding where they conflict with current practice or are confusing. For illustration, 1.4 l rather than 1.4 × 10-3 m3, or 4 mm somewhat than 4 × 10-3 m. Chemical formula and solutions must identify the form used, e.g. anhydrous or hydrated, and the concentration must be in clearly defined units. Common species names should be followed by underlines at the first mention. For following use the generic name should be constricted to a single letter, if it is clear.

Structure

All manuscripts submitted to Global Journals Inc. (US), ought to include:

Title: The title page must carry an instructive title that reflects the content, a running title (less than 45 characters together with spaces), names of the authors and co-authors, and the place(s) wherever the work was carried out. The full postal address in addition with the e-mail address of related author must be given. Up to eleven keywords or very brief phrases have to be given to help data retrieval, mining and indexing.

Abstract, used in Original Papers and Reviews:

Optimizing Abstract for Search Engines

Many researchers searching for information online will use search engines such as Google, Yahoo or similar. By optimizing your paper for search engines, you will amplify the chance of someone finding it. This in turn will make it more likely to be viewed and/or cited in a further work. Global Journals Inc. (US) have compiled these guidelines to facilitate you to maximize the web-friendliness of the most public part of your paper.

Key Words

A major linchpin in research work for the writing research paper is the keyword search, which one will employ to find both library and Internet resources.

One must be persistent and creative in using keywords. An effective keyword search requires a strategy and planning a list of possible keywords and phrases to try.

Search engines for most searches, use Boolean searching, which is somewhat different from Internet searches. The Boolean search uses "operators," words (and, or, not, and near) that enable you to expand or narrow your affords. Tips for research paper while preparing research paper are very helpful guideline of research paper.

Choice of key words is first tool of tips to write research paper. Research paper writing is an art.A few tips for deciding as strategically as possible about keyword search:

- One should start brainstorming lists of possible keywords before even begin searching. Think about the most important concepts related to research work. Ask, "What words would a source have to include to be truly valuable in research paper?" Then consider synonyms for the important words.
- It may take the discovery of only one relevant paper to let steer in the right keyword direction because in most databases, the keywords under which a research paper is abstracted are listed with the paper.
- One should avoid outdated words.

Keywords are the key that opens a door to research work sources. Keyword searching is an art in which researcher's skills are bound to improve with experience and time.

Numerical Methods: Numerical methods used should be clear and, where appropriate, supported by references.

Acknowledgements: Please make these as concise as possible.

References

References follow the Harvard scheme of referencing. References in the text should cite the authors' names followed by the time of their publication, unless there are three or more authors when simply the first author's name is quoted followed by et al. unpublished work has to only be cited where necessary, and only in the text. Copies of references in press in other journals have to be supplied with submitted typescripts. It is necessary that all citations and references be carefully checked before submission, as mistakes or omissions will cause delays.

References to information on the World Wide Web can be given, but only if the information is available without charge to readers on an official site. Wikipedia and Similar websites are not allowed where anyone can change the information. Authors will be asked to make available electronic copies of the cited information for inclusion on the Global Journals Inc. (US) homepage at the judgment of the Editorial Board.

The Editorial Board and Global Journals Inc. (US) recommend that, citation of online-published papers and other material should be done via a DOI (digital object identifier). If an author cites anything, which does not have a DOI, they run the risk of the cited material not being noticeable.

The Editorial Board and Global Journals Inc. (US) recommend the use of a tool such as Reference Manager for reference management and formatting.

Tables, Figures and Figure Legends

Tables: Tables should be few in number, cautiously designed, uncrowned, and include only essential data. Each must have an Arabic number, e.g. Table 4, a self-explanatory caption and be on a separate sheet. Vertical lines should not be used.

Figures: Figures are supposed to be submitted as separate files. Always take in a citation in the text for each figure using Arabic numbers, e.g. Fig. 4. Artwork must be submitted online in electronic form by e-mailing them.

Preparation of Electronic Figures for Publication

Even though low quality images are sufficient for review purposes, print publication requires high quality images to prevent the final product being blurred or fuzzy. Submit (or e-mail) EPS (line art) or TIFF (halftone/photographs) files only. MS PowerPoint and Word Graphics are unsuitable for printed pictures. Do not use pixel-oriented software. Scans (TIFF only) should have a resolution of at least 350 dpi (halftone) or 700 to 1100 dpi (line drawings) in relation to the imitation size. Please give the data for figures in black and white or submit a Color Work Agreement Form. EPS files must be saved with fonts embedded (and with a TIFF preview, if possible).

For scanned images, the scanning resolution (at final image size) ought to be as follows to ensure good reproduction: line art: >650 dpi; halftones (including gel photographs) : >350 dpi; figures containing both halftone and line images: >650 dpi.

Color Charges: It is the rule of the Global Journals Inc. (US) for authors to pay the full cost for the reproduction of their color artwork. Hence, please note that, if there is color artwork in your manuscript when it is accepted for publication, we would require you to complete and return a color work agreement form before your paper can be published.

Figure Legends: Self-explanatory legends of all figures should be incorporated separately under the heading 'Legends to Figures'. In the full-text online edition of the journal, figure legends may possibly be truncated in abbreviated links to the full screen version. Therefore, the first 100 characters of any legend should notify the reader, about the key aspects of the figure.

6. AFTER ACCEPTANCE

Upon approval of a paper for publication, the manuscript will be forwarded to the dean, who is responsible for the publication of the Global Journals Inc. (US).

6.1 Proof Corrections

The corresponding author will receive an e-mail alert containing a link to a website or will be attached. A working e-mail address must therefore be provided for the related author.

Acrobat Reader will be required in order to read this file. This software can be downloaded

(Free of charge) from the following website:

www.adobe.com/products/acrobat/readstep2.html. This will facilitate the file to be opened, read on screen, and printed out in order for any corrections to be added. Further instructions will be sent with the proof.

Proofs must be returned to the dean at dean@globaljournals.org within three days of receipt.

As changes to proofs are costly, we inquire that you only correct typesetting errors. All illustrations are retained by the publisher. Please note that the authors are responsible for all statements made in their work, including changes made by the copy editor.

6.2 Early View of Global Journals Inc. (US) (Publication Prior to Print)

The Global Journals Inc. (US) are enclosed by our publishing's Early View service. Early View articles are complete full-text articles sent in advance of their publication. Early View articles are absolute and final. They have been completely reviewed, revised and edited for publication, and the authors' final corrections have been incorporated. Because they are in final form, no changes can be made after sending them. The nature of Early View articles means that they do not yet have volume, issue or page numbers, so Early View articles cannot be cited in the conventional way.

6.3 Author Services

Online production tracking is available for your article through Author Services. Author Services enables authors to track their article - once it has been accepted - through the production process to publication online and in print. Authors can check the status of their articles online and choose to receive automated e-mails at key stages of production. The authors will receive an e-mail with a unique link that enables them to register and have their article automatically added to the system. Please ensure that a complete e-mail address is provided when submitting the manuscript.

6.4 Author Material Archive Policy

Please note that if not specifically requested, publisher will dispose off hardcopy & electronic information submitted, after the two months of publication. If you require the return of any information submitted, please inform the Editorial Board or dean as soon as possible.

6.5 Offprint and Extra Copies

A PDF offprint of the online-published article will be provided free of charge to the related author, and may be distributed according to the Publisher's terms and conditions. Additional paper offprint may be ordered by emailing us at: editor@globaljournals.org .

You must strictly follow above Author Guidelines before submitting your paper or else we will not at all be responsible for any corrections in future in any of the way.

Before start writing a good quality Computer Science Research Paper, let us first understand what is Computer Science Research Paper? So, Computer Science Research Paper is the paper which is written by professionals or scientists who are associated to Computer Science and Information Technology, or doing research study in these areas. If you are novel to this field then you can consult about this field from your supervisor or guide.

TECHNIQUES FOR WRITING A GOOD QUALITY RESEARCH PAPER:

1. Choosing the topic: In most cases, the topic is searched by the interest of author but it can be also suggested by the guides. You can have several topics and then you can judge that in which topic or subject you are finding yourself most comfortable. This can be done by asking several questions to yourself, like Will I be able to carry our search in this area? Will I find all necessary recourses to accomplish the search? Will I be able to find all information in this field area? If the answer of these types of questions will be "Yes" then you can choose that topic. In most of the cases, you may have to conduct the surveys and have to visit several places because this field is related to Computer Science and Information Technology. Also, you may have to do a lot of work to find all rise and falls regarding the various data of that subject. Sometimes, detailed information plays a vital role, instead of short information.

2. Evaluators are human: First thing to remember that evaluators are also human being. They are not only meant for rejecting a paper. They are here to evaluate your paper. So, present your Best.

3. Think Like Evaluators: If you are in a confusion or getting demotivated that your paper will be accepted by evaluators or not, then think and try to evaluate your paper like an Evaluator. Try to understand that what an evaluator wants in your research paper and automatically you will have your answer.

4. Make blueprints of paper: The outline is the plan or framework that will help you to arrange your thoughts. It will make your paper logical. But remember that all points of your outline must be related to the topic you have chosen.

5. Ask your Guides: If you are having any difficulty in your research, then do not hesitate to share your difficulty to your guide (if you have any). They will surely help you out and resolve your doubts. If you can't clarify what exactly you require for your work then ask the supervisor to help you with the alternative. He might also provide you the list of essential readings.

6. Use of computer is recommended: As you are doing research in the field of Computer Science, then this point is quite obvious.

7. Use right software: Always use good quality software packages. If you are not capable to judge good software then you can lose quality of your paper unknowingly. There are various software programs available to help you, which you can get through Internet.

8. Use the Internet for help: An excellent start for your paper can be by using the Google. It is an excellent search engine, where you can have your doubts resolved. You may also read some answers for the frequent question how to write my research paper or find model research paper. From the internet library you can download books. If you have all required books make important reading selecting and analyzing the specified information. Then put together research paper sketch out.

9. Use and get big pictures: Always use encyclopedias, Wikipedia to get pictures so that you can go into the depth.

10. Bookmarks are useful: When you read any book or magazine, you generally use bookmarks, right! It is a good habit, which helps to not to lose your continuity. You should always use bookmarks while searching on Internet also, which will make your search easier.

11. Revise what you wrote: When you write anything, always read it, summarize it and then finalize it.

12. Make all efforts: Make all efforts to mention what you are going to write in your paper. That means always have a good start. Try to mention everything in introduction, that what is the need of a particular research paper. Polish your work by good skill of writing and always give an evaluator, what he wants.

13. Have backups: When you are going to do any important thing like making research paper, you should always have backup copies of it either in your computer or in paper. This will help you to not to lose any of your important.

14. Produce good diagrams of your own: Always try to include good charts or diagrams in your paper to improve quality. Using several and unnecessary diagrams will degrade the quality of your paper by creating "hotchpotch." So always, try to make and include those diagrams, which are made by your own to improve readability and understandability of your paper.

15. Use of direct quotes: When you do research relevant to literature, history or current affairs then use of quotes become essential but if study is relevant to science then use of quotes is not preferable.

16. Use proper verb tense: Use proper verb tenses in your paper. Use past tense, to present those events that happened. Use present tense to indicate events that are going on. Use future tense to indicate future happening events. Use of improper and wrong tenses will confuse the evaluator. Avoid the sentences that are incomplete.

17. Never use online paper: If you are getting any paper on Internet, then never use it as your research paper because it might be possible that evaluator has already seen it or maybe it is outdated version.

18. Pick a good study spot: To do your research studies always try to pick a spot, which is quiet. Every spot is not for studies. Spot that suits you choose it and proceed further.

19. Know what you know: Always try to know, what you know by making objectives. Else, you will be confused and cannot achieve your target.

20. Use good quality grammar: Always use a good quality grammar and use words that will throw positive impact on evaluator. Use of good quality grammar does not mean to use tough words, that for each word the evaluator has to go through dictionary. Do not start sentence with a conjunction. Do not fragment sentences. Eliminate one-word sentences. Ignore passive voice. Do not ever use a big word when a diminutive one would suffice. Verbs have to be in agreement with their subjects. Prepositions are not expressions to finish sentences with. It is incorrect to ever divide an infinitive. Avoid clichés like the disease. Also, always shun irritating alliteration. Use language that is simple and straight forward. put together a neat summary.

21. Arrangement of information: Each section of the main body should start with an opening sentence and there should be a changeover at the end of the section. Give only valid and powerful arguments to your topic. You may also maintain your arguments with records.

22. Never start in last minute: Always start at right time and give enough time to research work. Leaving everything to the last minute will degrade your paper and spoil your work.

23. Multitasking in research is not good: Doing several things at the same time proves bad habit in case of research activity. Research is an area, where everything has a particular time slot. Divide your research work in parts and do particular part in particular time slot.

24. Never copy others' work: Never copy others' work and give it your name because if evaluator has seen it anywhere you will be in trouble.

25. Take proper rest and food: No matter how many hours you spend for your research activity, if you are not taking care of your health then all your efforts will be in vain. For a quality research, study is must, and this can be done by taking proper rest and food.

26. Go for seminars: Attend seminars if the topic is relevant to your research area. Utilize all your resources.

27. Refresh your mind after intervals: Try to give rest to your mind by listening to soft music or by sleeping in intervals. This will also improve your memory.

28. Make colleagues: Always try to make colleagues. No matter how sharper or intelligent you are, if you make colleagues you can have several ideas, which will be helpful for your research.

29. Think technically: Always think technically. If anything happens, then search its reasons, its benefits, and demerits.

30. Think and then print: When you will go to print your paper, notice that tables are not be split, headings are not detached from their descriptions, and page sequence is maintained.

31. Adding unnecessary information: Do not add unnecessary information, like, I have used MS Excel to draw graph. Do not add irrelevant and inappropriate material. These all will create superfluous. Foreign terminology and phrases are not apropos. One should NEVER take a broad view. Analogy in script is like feathers on a snake. Not at all use a large word when a very small one would be sufficient. Use words properly, regardless of how others use them. Remove quotations. Puns are for kids, not grunt readers. Amplification is a billion times of inferior quality than sarcasm.

32. Never oversimplify everything: To add material in your research paper, never go for oversimplification. This will definitely irritate the evaluator. Be more or less specific. Also too, by no means, ever use rhythmic redundancies. Contractions aren't essential and shouldn't be there used. Comparisons are as terrible as clichés. Give up ampersands and abbreviations, and so on. Remove commas, that are, not necessary. Parenthetical words however should be together with this in commas. Understatement is all the time the complete best way to put onward earth-shaking thoughts. Give a detailed literary review.

33. Report concluded results: Use concluded results. From raw data, filter the results and then conclude your studies based on measurements and observations taken. Significant figures and appropriate number of decimal places should be used. Parenthetical remarks are prohibitive. Proofread carefully at final stage. In the end give outline to your arguments. Spot out perspectives of further study of this subject. Justify your conclusion by at the bottom of them with sufficient justifications and examples.

34. After conclusion: Once you have concluded your research, the next most important step is to present your findings. Presentation is extremely important as it is the definite medium though which your research is going to be in print to the rest of the crowd. Care should be taken to categorize your thoughts well and present them in a logical and neat manner. A good quality research paper format is essential because it serves to highlight your research paper and bring to light all necessary aspects in your research.

Informal Guidelines of Research Paper Writing

Key points to remember:

- Submit all work in its final form.
- Write your paper in the form, which is presented in the guidelines using the template.
- Please note the criterion for grading the final paper by peer-reviewers.

Final Points:

A purpose of organizing a research paper is to let people to interpret your effort selectively. The journal requires the following sections, submitted in the order listed, each section to start on a new page.

The introduction will be compiled from reference matter and will reflect the design processes or outline of basis that direct you to make study. As you will carry out the process of study, the method and process section will be constructed as like that. The result segment will show related statistics in nearly sequential order and will direct the reviewers next to the similar intellectual paths throughout the data that you took to carry out your study. The discussion section will provide understanding of the data and projections as to the implication of the results. The use of good quality references all through the paper will give the effort trustworthiness by representing an alertness of prior workings.

Writing a research paper is not an easy job no matter how trouble-free the actual research or concept. Practice, excellent preparation, and controlled record keeping are the only means to make straightforward the progression.

General style:

Specific editorial column necessities for compliance of a manuscript will always take over from directions in these general guidelines.

To make a paper clear

· Adhere to recommended page limits

Mistakes to evade

- Insertion a title at the foot of a page with the subsequent text on the next page
- Separating a table/chart or figure - impound each figure/table to a single page
- Submitting a manuscript with pages out of sequence

In every sections of your document

· Use standard writing style including articles ("a", "the," etc.)

· Keep on paying attention on the research topic of the paper

· Use paragraphs to split each significant point (excluding for the abstract)

· Align the primary line of each section

· Present your points in sound order

· Use present tense to report well accepted

· Use past tense to describe specific results

· Shun familiar wording, don't address the reviewer directly, and don't use slang, slang language, or superlatives

· Shun use of extra pictures - include only those figures essential to presenting results

Title Page:

Choose a revealing title. It should be short. It should not have non-standard acronyms or abbreviations. It should not exceed two printed lines. It should include the name(s) and address (es) of all authors.

Abstract:

The summary should be two hundred words or less. It should briefly and clearly explain the key findings reported in the manuscript--must have precise statistics. It should not have abnormal acronyms or abbreviations. It should be logical in itself. Shun citing references at this point.

An abstract is a brief distinct paragraph summary of finished work or work in development. In a minute or less a reviewer can be taught the foundation behind the study, common approach to the problem, relevant results, and significant conclusions or new questions.

Write your summary when your paper is completed because how can you write the summary of anything which is not yet written? Wealth of terminology is very essential in abstract. Yet, use comprehensive sentences and do not let go readability for briefness. You can maintain it succinct by phrasing sentences so that they provide more than lone rationale. The author can at this moment go straight to shortening the outcome. Sum up the study, with the subsequent elements in any summary. Try to maintain the initial two items to no more than one ruling each.

- Reason of the study - theory, overall issue, purpose
- Fundamental goal
- To the point depiction of the research
- Consequences, including definite statistics - if the consequences are quantitative in nature, account quantitative data; results of any numerical analysis should be reported
- Significant conclusions or questions that track from the research(es)

Approach:

- Single section, and succinct
- As a outline of job done, it is always written in past tense
- A conceptual should situate on its own, and not submit to any other part of the paper such as a form or table
- Center on shortening results - bound background information to a verdict or two, if completely necessary
- What you account in an conceptual must be regular with what you reported in the manuscript
- Exact spelling, clearness of sentences and phrases, and appropriate reporting of quantities (proper units, important statistics) are just as significant in an abstract as they are anywhere else

Introduction:

The **Introduction** should "introduce" the manuscript. The reviewer should be presented with sufficient background information to be capable to comprehend and calculate the purpose of your study without having to submit to other works. The basis for the study should be offered. Give most important references but shun difficult to make a comprehensive appraisal of the topic. In the introduction, describe the problem visibly. If the problem is not acknowledged in a logical, reasonable way, the reviewer will have no attention in your result. Speak in common terms about techniques used to explain the problem, if needed, but do not present any particulars about the protocols here. Following approach can create a valuable beginning:

- Explain the value (significance) of the study
- Shield the model - why did you employ this particular system or method? What is its compensation? You strength remark on its appropriateness from a abstract point of vision as well as point out sensible reasons for using it.
- Present a justification. Status your particular theory (es) or aim(s), and describe the logic that led you to choose them.
- Very for a short time explain the tentative propose and how it skilled the declared objectives.

Approach:

- Use past tense except for when referring to recognized facts. After all, the manuscript will be submitted after the entire job is done.
- Sort out your thoughts; manufacture one key point with every section. If you make the four points listed above, you will need a least of four paragraphs.

- Present surroundings information only as desirable in order hold up a situation. The reviewer does not desire to read the whole thing you know about a topic.
- Shape the theory/purpose specifically - do not take a broad view.
- As always, give awareness to spelling, simplicity and correctness of sentences and phrases.

Procedures (Methods and Materials):

This part is supposed to be the easiest to carve if you have good skills. A sound written Procedures segment allows a capable scientist to replacement your results. Present precise information about your supplies. The suppliers and clarity of reagents can be helpful bits of information. Present methods in sequential order but linked methodologies can be grouped as a segment. Be concise when relating the protocols. Attempt for the least amount of information that would permit another capable scientist to spare your outcome but be cautious that vital information is integrated. The use of subheadings is suggested and ought to be synchronized with the results section. When a technique is used that has been well described in another object, mention the specific item describing a way but draw the basic principle while stating the situation. The purpose is to text all particular resources and broad procedures, so that another person may use some or all of the methods in one more study or referee the scientific value of your work. It is not to be a step by step report of the whole thing you did, nor is a methods section a set of orders.

Materials:

- Explain materials individually only if the study is so complex that it saves liberty this way.
- Embrace particular materials, and any tools or provisions that are not frequently found in laboratories.
- Do not take in frequently found.
- If use of a definite type of tools.
- Materials may be reported in a part section or else they may be recognized along with your measures.

Methods:

- Report the method (not particulars of each process that engaged the same methodology)
- Describe the method entirely
- To be succinct, present methods under headings dedicated to specific dealings or groups of measures
- Simplify - details how procedures were completed not how they were exclusively performed on a particular day.
- If well known procedures were used, account the procedure by name, possibly with reference, and that's all.

Approach:

- It is embarrassed or not possible to use vigorous voice when documenting methods with no using first person, which would focus the reviewer's interest on the researcher rather than the job. As a result when script up the methods most authors use third person passive voice.
- Use standard style in this and in every other part of the paper - avoid familiar lists, and use full sentences.

What to keep away from

- Resources and methods are not a set of information.
- Skip all descriptive information and surroundings - save it for the argument.
- Leave out information that is immaterial to a third party.

Results:

The principle of a results segment is to present and demonstrate your conclusion. Create this part a entirely objective details of the outcome, and save all understanding for the discussion.

The page length of this segment is set by the sum and types of data to be reported. Carry on to be to the point, by means of statistics and tables, if suitable, to present consequences most efficiently.You must obviously differentiate material that would usually be incorporated in a study editorial from any unprocessed data or additional appendix matter that would not be available. In fact, such matter should not be submitted at all except requested by the instructor.

Content

- Sum up your conclusion in text and demonstrate them, if suitable, with figures and tables.
- In manuscript, explain each of your consequences, point the reader to remarks that are most appropriate.
- Present a background, such as by describing the question that was addressed by creation an exacting study.
- Explain results of control experiments and comprise remarks that are not accessible in a prescribed figure or table, if appropriate.
- Examine your data, then prepare the analyzed (transformed) data in the form of a figure (graph), table, or in manuscript form.

What to stay away from

- Do not discuss or infer your outcome, report surroundings information, or try to explain anything.
- Not at all, take in raw data or intermediate calculations in a research manuscript.
- Do not present the similar data more than once.
- Manuscript should complement any figures or tables, not duplicate the identical information.
- Never confuse figures with tables - there is a difference.

Approach

- As forever, use past tense when you submit to your results, and put the whole thing in a reasonable order.
- Put figures and tables, appropriately numbered, in order at the end of the report
- If you desire, you may place your figures and tables properly within the text of your results part.

Figures and tables

- If you put figures and tables at the end of the details, make certain that they are visibly distinguished from any attach appendix materials, such as raw facts
- Despite of position, each figure must be numbered one after the other and complete with subtitle
- In spite of position, each table must be titled, numbered one after the other and complete with heading
- All figure and table must be adequately complete that it could situate on its own, divide from text

Discussion:

The Discussion is expected the trickiest segment to write and describe. A lot of papers submitted for journal are discarded based on problems with the Discussion. There is no head of state for how long a argument should be. Position your understanding of the outcome visibly to lead the reviewer through your conclusions, and then finish the paper with a summing up of the implication of the study. The purpose here is to offer an understanding of your results and hold up for all of your conclusions, using facts from your research and generally accepted information, if suitable. The implication of result should be visibly described. Infer your data in the conversation in suitable depth. This means that when you clarify an observable fact you must explain mechanisms that may account for the observation. If your results vary from your prospect, make clear why that may have happened. If your results agree, then explain the theory that the proof supported. It is never suitable to just state that the data approved with prospect, and let it drop at that.

- Make a decision if each premise is supported, discarded, or if you cannot make a conclusion with assurance. Do not just dismiss a study or part of a study as "uncertain."
- Research papers are not acknowledged if the work is imperfect. Draw what conclusions you can based upon the results that you have, and take care of the study as a finished work
- You may propose future guidelines, such as how the experiment might be personalized to accomplish a new idea.
- Give details all of your remarks as much as possible, focus on mechanisms.
- Make a decision if the tentative design sufficiently addressed the theory, and whether or not it was correctly restricted.
- Try to present substitute explanations if sensible alternatives be present.
- One research will not counter an overall question, so maintain the large picture in mind, where do you go next? The best studies unlock new avenues of study. What questions remain?
- Recommendations for detailed papers will offer supplementary suggestions.

Approach:

- When you refer to information, differentiate data generated by your own studies from available information
- Submit to work done by specific persons (including you) in past tense.
- Submit to generally acknowledged facts and main beliefs in present tense.

Please carefully note down following rules and regulation before submitting your Research Paper to Global Journals Inc. (US):

Segment Draft and Final Research Paper: You have to strictly follow the template of research paper. If it is not done your paper may get rejected.

- The **major constraint** is that you must independently make all content, tables, graphs, and facts that are offered in the paper. You must write each part of the paper wholly on your own. The Peer-reviewers need to identify your own perceptive of the concepts in your own terms. NEVER extract straight from any foundation, and never rephrase someone else's analysis.

- Do not give permission to anyone else to "PROOFREAD" your manuscript.

- Methods to avoid Plagiarism is applied by us on every paper, if found guilty, you will be blacklisted by all of our collaborated research groups, your institution will be informed for this and strict legal actions will be taken immediately.)
- To guard yourself and others from possible illegal use please do not permit anyone right to use to your paper and files.

Criterion for Grading a Research Paper (Compilation) by Global Journals Inc. (US)

Please note that following table is only a Grading of "Paper Compilation" and not on "Performed/Stated Research" whose grading solely depends on Individual Assigned Peer Reviewer and Editorial Board Member. These can be available only on request and after decision of Paper. This report will be the property of Global Journals Inc. (US).

Topics	Grades		
	A-B	**C-D**	**E-F**
Abstract	Clear and concise with appropriate content, Correct format. 200 words or below	Unclear summary and no specific data, Incorrect form Above 200 words	No specific data with ambiguous information Above 250 words
Introduction	Containing all background details with clear goal and appropriate details, flow specification, no grammar and spelling mistake, well organized sentence and paragraph, reference cited	Unclear and confusing data, appropriate format, grammar and spelling errors with unorganized matter	Out of place depth and content, hazy format
Methods and Procedures	Clear and to the point with well arranged paragraph, precision and accuracy of facts and figures, well organized subheads	Difficult to comprehend with embarrassed text, too much explanation but completed	Incorrect and unorganized structure with hazy meaning
Result	Well organized, Clear and specific, Correct units with precision, correct data, well structuring of paragraph, no grammar and spelling mistake	Complete and embarrassed text, difficult to comprehend	Irregular format with wrong facts and figures
Discussion	Well organized, meaningful specification, sound conclusion, logical and concise explanation, highly structured paragraph reference cited	Wordy, unclear conclusion, spurious	Conclusion is not cited, unorganized, difficult to comprehend
References	Complete and correct format, well organized	Beside the point, Incomplete	Wrong format and structuring

INDEX

save our planet

Global Journal of Computer Science and Technology

Visit us on the Web at www.GlobalJournals.org | www.ComputerResearch.org
or email us at helpdesk@globaljournals.org

ISSN 9754350