Online ISSN : 0975-4172
Print ISSN : 0975-4350

Global Journal of Computer Science & Technology

DISCOVERING THOUGHTS AND INVENTING FUTURE

Highlights

Making in Textile Sector

Tree Based Authenticated

Expression Recognition System

Length Frequent Sequences

Laboratory, Antarctica

Volume 12 | Issue 7 | Version 1.0

ENG

Global Journal of Computer Science and Technology

Global Journal of Computer Science and Technology

Volume 12 Issue 7 (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:Global Association of Research

Open Scientific Standards

Publisher's Headquarters office

Global Journals Inc., Headquarters Corporate Office, Cambridge Office Center, II Canal Park, Floor No. 5th, ***Cambridge (Massachusetts)***, Pin: MA 02141 United States

USA Toll Free: +001-888-839-7392

USA Toll Free Fax: +001-888-839-7392

Offset Typesetting

Global Association of Research, Marsh Road, Rainham, Essex, London RM13 8EU United Kingdom.

Packaging & Continental Dispatching

Global Journals, 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: *investers@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)

Editorial Board Members (HON.)

Dr. Bart Lambrecht
Director of Research in Accounting and FinanceProfessor 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 ReginaPh.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 FinanceLancaster University Management SchoolPh.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 NeuroscienceNorthwestern 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 SciencesPurdue University Director
National Center for Typhoon and Flooding Research, Taiwan
University Chair Professor
Department of Atmospheric Sciences, National Central University, Chung-Li, TaiwanUniversity 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 Volume

Global Journal of Computer Science and Technology
Volume 12 Issue 7 Version 1.0 April 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Usability of Data Warehousing and Data Mining for Interactive Decision Making in Textile Sector

By Muhammad Shakeel Faridi & Tasleem Mustafa
University of Agriculture, Faisalabad, Pakistan

Abstract - Data warehouse is one of the most rapidly growing areas in management information system. With this approach, data for Executive Information System (EIS) and Decision Support System (DSS) applications are separated from operational data and stored in a separate database. This process is called data warehousing. The major advantages of this approach are: improved in performance, better data quality, and the ability to consolidate and summarize data from heterogeneous systems. A data warehouse is part of a larger infrastructure that includes legacy data sources, external data sources, a repository, data acquisition software, and user interface and related analytical tools. The aim of this research work is to elaborate that how the textile industry can manage and improve their production capacity and resources at optimum level to produce a good quality result using data warehousing and data mining techniques. This research work is conduction in Masood Textile Mills Limited, Faisalabad, Pakistan (MTML). The results may hopefully opened-up an era of research and methodology that could further benefit the Industry to support in decision support system.

Keywords : *Executive Information System, Decision Support System, data warehouse, data mining, production capacity, heterogeneous systems.*

GJCST Classification: *H.2.8*

Strictly as per the compliance and regulations of:

Usability of Data Warehousing and Data Mining for Interactive Decision Making in Textile Sector

Muhammad Shakeel Faridi [α] & Tasleem Mustafa [σ]

Abstract - Data warehouse is one of the most rapidly growing areas in management information system. With this approach, data for Executive Information System (EIS) and Decision Support System (DSS) applications are separated from operational data and stored in a separate database. This process is called data warehousing. The major advantages of this approach are: improved in performance, better data quality, and the ability to consolidate and summarize data from heterogeneous systems. A data warehouse is part of a larger infrastructure that includes legacy data sources, external data sources, a repository, data acquisition software, and user interface and related analytical tools. The aim of this research work is to elaborate that how the textile industry can manage and improve their production capacity and resources at optimum level to produce a good quality result using data warehousing and data mining techniques. This research work is conduction in Masood Textile Mills Limited, Faisalabad, Pakistan (MTML). The results may hopefully opened-up an era of research and methodology that could further benefit the Industry to support in decision support system.

Keywords : Executive Information System, Decision Support System, data warehouse, data mining, production capacity, heterogeneous systems.

I. Introduction

Data warehousing is used to support in decision making. It collects data from multiple heterogeneous operational source systems, storing summarized integrated business information in a central repository used by analytical application with different user perspective. It is a process of merging data in a centralized location. The data warehouses' databases usually store the complete organizational business history from start to end. The frequently used procedure for fetching decision making information from data warehouses is based on executing queries and tools which perform online analytical processing (OLAP) [1]. Data mining is a rising methodology used for identifying, extracting and evaluating variables to find useful information [2]. Data Mining is a process of extracting previously stored, valid, potential by useful and hidden patterns from large data sets as data volume is increasing rapidly day by day [3]. It allow the users to analyze data from different dimensions categorize and summarize the relationships, identified during the mining process [4]. Different data mining techniques are used in various organizations and institutions like pharmaceutical, telecommunication, engineering, education, banking, marketing, sale, etc [5]. It is very difficult to normalize data without using any sort of technique. To get required benefits and useful information from such large amount of stored data, you must use some data mining techniques like Classical Techniques: Statistics, Neighborhoods and Clustering and Next Generation Techniques: Trees, Networks and Rules [4]. These sorts of techniques facilitate the user to get useful and accurate information on time for interactive decision.

The objective of this study is to elaborate that how the textile industries can manage and improve their production capacity and resources at optimum level to produce good quality results comparing from the historic information using data warehousing and data mining techniques.

II. The Usability Concept

Usability is defined as the capability in human functional terms that to be used easily, effectively and satisfactorily by specific users, performing specific responsibilities, in specific environments [23]. It can be measured analytically through the use of inspection method [7; 8]. The analytical approach to measuring usability attempts to identify actual usability issues that are practically faced by the end users [7; 9]. Usability inspection is regarded as a cost-effective approach (to measuring usability) which relies on the review of usability experts [7]. Inspection techniques differ from empirical techniques by identifying potential usability issues, as opposed to actual usability issues.

The definition of usability or usability engineering by international standardization organization (ISO) and usability-experts also defined it as "the effectiveness, efficiency and satisfaction with which specified user can achieve the specified goals in particular environment" [10, 11, 12, 13, 14, 15]. The basic purpose of usability is to make sure that the work is smoothly going done by the end user by using his average ability without any frustration [16, 27]. It not only discusses but also applies to the design of software system, applications and product but also includes users' interface, supporting documents to run the software efficiently [18, 19]. There is surety for un-successfulness of a fully functional software or application, if it has poor usability. As functionality and

Author α : Dept of Computer Science, University of Agriculture, Faisalabad, 38000, Pakistan. E-mail : shakeelfaridi@gmail.com
Author σ : Dept of Computer Science, University of Agriculture, Faisalabad, 38000, Pakistan. E-mail : tasleem_mustafa@uaf.edu.pk

usability both are tasks and user related terms [20]. Functions need to match task requirements while the users need to understand the exact functionality of the system to meet their requirements [21]. "Grouping related commands into menus, other examples of good interface design and documentation help overcome the conflict between power and ease of use to enhance both usability and functionality [22]." Evaluator's experience, available resources like time, budget, and labor are the multiple methods to evaluate the usability. It also depends on the stage of development tools.

III. Common Usability Evaluation Principles

There are some common usability evaluations principles are used to evaluate the requirements. These evaluation principles are; ease of use, usefulness, customization, task support, flexibility, navigation, guidance, memorability, system reliability, user interface, output presentation, learn ability, system responsiveness, accuracy and completeness [6, 24, 25].

The usability evaluation principles like usefulness, guidance, customization, accuracy and completeness are showing the amazing results in my work.

IV. How to Improve Decision Making by Using Data Mining

Hence, data mining uses predictive techniques to expose patterns in the data. These patterns of a data playing a vital role in the process of decision making since they expose the areas where improvements is required in the whole process. Organizations can use data mining in such a way as to improve profitableness and effectiveness of their interactions with their customers, improve the management of risk, and detect fraud. In other words, these patterns that are expose by using data mining techniques to assist business organizations and stakeholders of the organizations to make timely, accurately and better decisions [26].

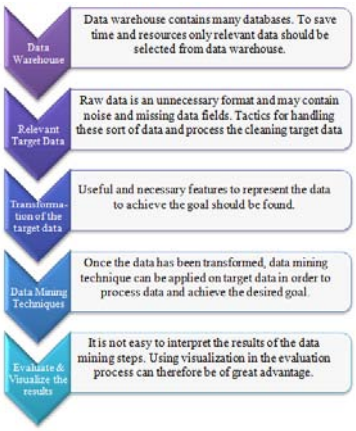


Figure 1: Data warehouse to Data Mining Process

a) Data Mining Techniques

There are some most popular data mining techniques: Prediction, Estimation, Regression and Classification.

i. *Prediction*

Forecasting, like estimation, is everywhere in business. Accurate prediction can reduce the cost, optimize your sales and better utilization of available resources. If you can predict the future, you will definitely boost your business and bring more profits [17].

ii. *Estimation*

This process is useful in just about every surface of business. From finance to marketing to sales, the better you can estimate your expenses, product mix optimization, or potential customer value, the better off you will be. This and the next use are fairly self-evident if you have ever spent a day at a business [17].

iii. *Regression*

This is the most widely known and the oldest statistical technique that is utilized by the data mining community. Essentially, regression makes use of a dataset to develop a mathematical formula which fits the data. So whenever you want to use the results for predicting future behavioral patterns, all you need to do is just take the new data, and apply it to the model that has been developed, and you will get your prediction. But if you need to work with data that is categorical, where there is no significant order, such as gender, name, or color, it is better to use a different technique.[26]

iv. *Classification*

If you need to work with categorical data, or a combination of categorical data and continuous numeric, classification analysis will meet your requirements. This technique has the capability to process a more extensive variety of data compared to regression and is therefore increasingly popular. The complex mathematical formula that the regression technique provides, in this you will be provided a decision tree which requires a sequence of binary decisions. [26]

V. Empirical Work

Computer and related technologies are playing a vital role in any organization. The proper usage and implementation of these technologies have made the organizations to rule in the world. MTM (Masood Textile Mills Limited) is a top manufacturing unit having inhouse spinning, knitting, dyeing and garments facilitates. MTM have full capability of utilization of information technology in all areas. It has highly professional and technical staff engaged at every technical and managerial stage. All key professionals are involved in decision making. Delegation of power is made with

proper accountability. All production operations are fully computerized starting from start to end of products. This makes the MTM one of the best apparel manufacturer in Pakistan.

MTM have their complete and fully in-house developed ERP system. It contains all important applications that provide the absolute solutions for its operations and management as well. This well designed ERP system makes MTM competitive advantages over the others in this region.

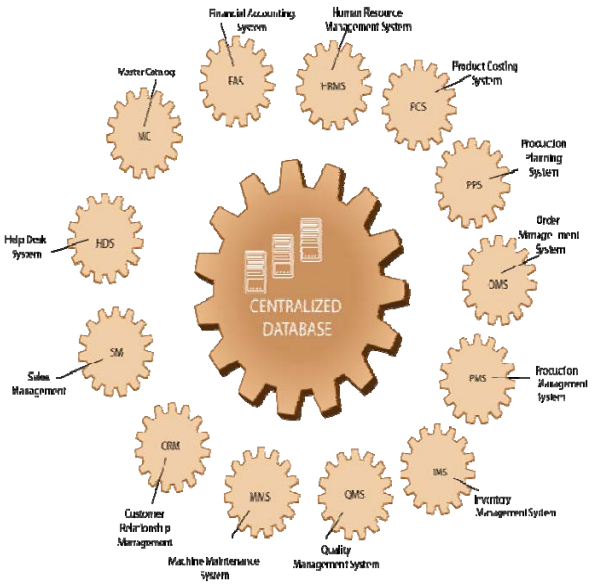


Figure 2 : Centralized Database

Figure 2 shows that the centralized database is connected with different application modules of the organizations. These application modules are located at different places which are connected to centralized database. The production server incorporate huge amount of data across the multiple sites of the organization. This facilitates Company-wide integrated Information Systems Covering all functional Areas. The top management of the organization has capability to monitor the whole areas of business process. They can control, monitor and apply the business rules by using online application modules.

VI. Result and Analysis

The top management of MTM like CEO, Director Operations has fully focused on capacity utilization to boost up the organization. The given data in Table 1 is collected from different data bases to present it to the management of MTM in summarized form. This data shows the available and utilized knitting capacity on yearly basis.

The data gathered from backup servers and production servers to compare the yearly analysis by using data mining techniques. The actual capacity is 15,455, 19,163, 19,216, 19,654, 25,848 and 28,270 from year 2006 to 2011 while the utilization of capacity is 6,491, 7,910, 10,211, 10,458, 14,165, 14,484 from year 2006 to 2011 which shows that the utilized capacity from year 2006 to 2011 is 42%, 41%, 53%, 53%, 55%, and 51% where almost 50 percent capacity losses are borne by the company on yearly basis and at what capacity the industry has been working over the last six years.

Figures in ('000')

Year	Actual Capacity	Capacity Utilized	% age Utilized
2006	15,455	6,491	42
2007	19,163	7,910	41
2008	19,216	10,211	53
2009	19,654	10,458	53
2010	25,848	14,165	55
2011	28,270	14,484	51

Table 1: knitting Production Capacity & Utilization

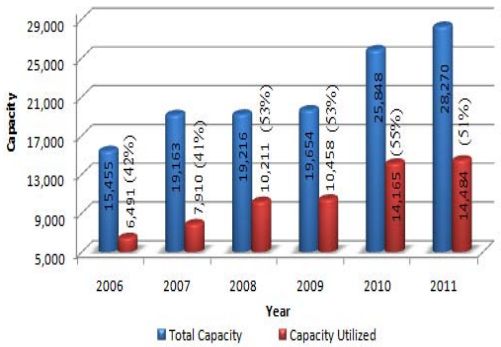


Figure 3 : Knitting Production Actual vs Utilized Figures in ('000')

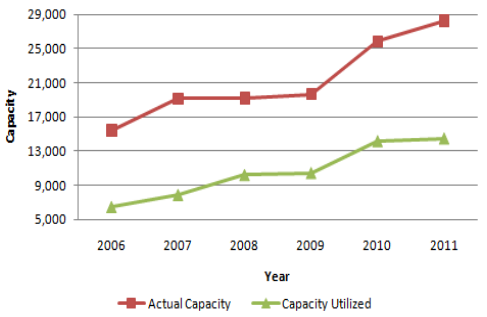


Figure 4 : Knitting Production Actual vs Utilized Figures in ('000')

Figure 3 and *Figure 4* are graphically shows the actual and utilized capacity.

VII. Conclusion

The value of data warehousing and data mining can be seen in effective decision making based on concrete evidence from the old data. The purpose of this research was basically to highlight the core area of business issues. These core areas, where top management needs to be focused to make the organization more profitable by using latest technologies. In order to continue aggressive data mining techniques must be used to save money, achieve better operational performance and mitigate risk. Cost saving will be achieved by maintaining or improving production performance. It would be difficult to achieve more profitability by ignoring these loop holes. The present study was focused only one area of the business by using data mining for decision analysis. In future the author will try to discuss the other core area of business that can directly affect the organization's objectives.

References Références Referencias

1. S. Chaudhuri, and U. Dayal: "An overview of data warehousing and OLAP Technology", SIGMOD Record, Volume 26, Number 1, PP. 65-74, 1997.
2. Alaa el-Halees (2009) Mining Students Data to Analyze e-Learning Behavior: A Case Study.
3. Connolly T., C. Begg and A. Strachan (1999) Database Systems: A Practical Approach to Design, Implementation, and Management (3rd Ed.). Harlow: Addison-Wesley.687
4. Han,J. and Kamber, M., (2006) "Data Mining: Concepts and Techniques", 2nd edition. The Morgan Kaufmann Series in Data Management Systems, Jim Gray, Series Editor.
5. ZhaoHui. Maclennan.J, (2005). Data Mining with SQL Server 2005 Wihely Publishing, Inc.
6. Herbert, L., Ragsdale, J. and Gaynor, E. (2006): Put Business Applications to the Usability Test. Forrester Research.
7. Hollingsed, T. and Novick, D.G. (2007): Usability inspection methods after 15 years of research and practice. In Proceedings of The 25th Annual ACM international Conference on Design of Communication, El Paso, Texas, USA. 249-255, ACM Press
8. Hornbæk, K. (2006): Current practice in measuring usability: Challenges to usability studies and research. International Journal of Human-Computer Studies 64(2):79-102
9. Schmettow, M. and Vietz, W. (2008): Introducing Item Response Theory for Measuring Usability Inspection Processes. In Proceedings of The Twenty- Sixth Annual SIGCHI Conference on Human Factors in Computing Systems, Florance, Italy. 893-902, ACM Press
10. P. W. Jordan, An introduction to usability, Taylor and Francis Ltd, 1998, pp.1-7.
11. D. Benyon, Phil Tuner and Susan Tuner, Designing the interactive system. Harlow: Person education limited, 2005, pp.5-6.
12. Preece et al., Interaction design, New York: jhon Wiley & sons, 2002, pp. 14, 230, 390-398.
13. J. Nielsen, Usability Engineering. Academic Press, 1993, pp. 25-26, 196-214.
14. X. Faulkner, Usability Engineering. Macmillan press Ltd, 2000, pp. 7.
15. T. Jokela, N. Iivari, "The Standard of User-Centered Design and the Standard Definition of Usability: Analyzing ISO 13407 against ISO9241-11" Proceedings of the Latin American conference on Human-computer interaction; Vol. 46, pp.53-60, 2003.
16. Krug, S. Don't Make Me Think: A Common Sens Approach to Web Usability. New Riders Publishing,Indianapolis, IN, 2000.
17. Business Intelligence data mining techniques: Available at http://www.chicagobigroup.com/ business_intelli-gence_white_papers/business_inte-lligence_data_mining_techniques.pdf [Accessed on December, 2011] Norman, D. A. The Psychology of Everyday Things Basic Books, NY, (1988).
18. Hohmann, L. Usability: Happier users mean greater profits. Information Systems Management, 20, 4, 66- 76, 2003.
19. Davis, F.D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13, 319 340, 1989.
20. Muhammad Shakeel Faridi, "Usability Evaluation of In-Housed Developed ERP System", http://spiedigitallibrary.org/proceedings/resource/2/ psis dg/8285/1/82850C_1?isAuthorized=no Proc. SPIE 8285, 82850C (2011); doi:10.1117/12.913212
21. Technology Acceptance and ERP Documentation Usability by Judy E. Scott. doi: 10. 1145/1400214. 1400239
22. Shackel, B. (1991) Usability - Context, Framework, Definition, Design and Evaluation. In B. Shackel and S. Richardson (eds.) Human Factors for Informatics Usability (Cambridge: Cambridge University Press) 21-38.
23. Calisir, F. and Calisir, F. (2004): The relation of interface usability characteristics, perceived usefulness, and perceived ease of use to end-user satisfaction with enterprise resource planning (ERP) systems. Computers in Behavior 20
24. Sap (2004): User Day Toolkit [online]. Available at http://www.sapdesignguild.org/resources/user_day_ tool kit/index.htm [Accessed on March]
25. Data mining techniques and applications: Available at http://www.buzzle.com/articles/data-mining tech-

niques- and-applications.html [Accessed on December, 2011]
26. Steve Krug, Don't Make Me Think, 2000, p. 5

Global Journal of Computer Science and Technology Volume XII Issue VII Version I April 2012
6

This page is intentionally left blank

Global Journal of Computer Science and Technology
Volume 12 Issue 7 Version 1.0 April 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

A Neural Network Based Classifier for a Segmented Facial Expression Recognition System Based on Haar Wavelet Transform

By Ongalo P. N. Fedha, Huang Dong Jun & Richard Rimiru
Central South University, Changsha, China

Abstract - Automatic recognition of facial expressions is a vital component of natural human-machine interfaces. Facial expressions convey information about one's emotional state and helps regulate our social norms by helping detect and interpret a scene. In this paper, we propose a novel face expression recognition scheme based on Haar discrete wavelet transform and a neural network classifier. First, the sample image undergoes preprocessing where noise is removed using binary image processing techniques. Then feature vectors are extracted using DWT from corresponding pixels in the image. The extracted image pixel data are used as the input to the neural network. We demonstrate experimentally that when wavelet coefficients are fed into a back-propagation based neural network for classification, a high recognition rate can be achieved by using a very small proportion of transform coefficients. Based on our experimental results, the proposed scheme gives satisfactory results.

Keywords : *Discrete wavelet transform, Neural network classifier, Facial expression recognition.*

GJCST Classification: *I.2.6*

Strictly as per the compliance and regulations of:

A Neural Network Based Classifier for a Segmented Facial Expression Recognition System Based on Haar Wavelet Transform

Ongalo P. N. Fedha [α], Huang Dong Jun [σ] & Richard Rimiru [ρ]

Abstract - Automatic recognition of facial expressions is a vital component of natural human-machine interfaces. Facial expressions convey information about one's emotional state and helps regulate our social norms by helping detect and interpret a scene. In this paper, we propose a novel face expression recognition scheme based on Haar discrete wavelet transform and a neural network classifier. First, the sample image undergoes preprocessing where noise is removed using binary image processing techniques. Then feature vectors are extracted using DWT from corresponding pixels in the image. The extracted image pixel data are used as the input to the neural network. We demonstrate experimentally that when wavelet coefficients are fed into a back-propagation based neural network for classification, a high recognition rate can be achieved by using a very small proportion of transform coefficients. Based on our experimental results, the proposed scheme gives satisfactory results.

Keywords : Discrete wavelet transform, Neural network classifier, Facial expression recognition.

I. Introduction

The human face is undoubtedly the most common characteristic used by humans to recognize and reflect face expressions with speed and accuracy in a passive and non intrusive manner. Facial expressions deals with the classification of facial muscles motion and facial feature deformation into abstract classes that are purely based on visual information while human emotions result from other different factors and their state might or might not be revealed through a number of communication channels such as emotional voice, pose, gestures, gaze direction and facial expressions [1]. Consider a scenario where a person tries downloading one of his favorite movies but becomes frustrated with the system's inability to load the program due to bandwidth or some other reason. In this case the person is likely to express some sort of emotional dissatisfaction well captured via face expressions either intentionally or unintentionally. There are existing works like Chen et al. [2] and De Silva et al.[3] that studied combined detection of facial and vocal expressions of emotion. However, majority of the studies treat the various human communication channels separately.

An automatic recognition of face expression through facial Action Unit (AU) has attracted much attention in the recent years due to its potential applications in behavioral science, medicine, security and human machine interaction. The Facial Action Coding System (FACS) developed by Ekman and Friesen [4] is the most commonly used system for facial behavior analysis. Based on image data, face expressions can be categorized based on whether they are static images or dynamic image sequence. Static images analyze single, still image, based on the spatial information of a frame and the face geometry. It has less computation and is more suitable for real-time facial expression recognition. Dynamic image sequences on the other hand take the motion information of expression images, the expression changes in time and space information together into account, so the recognition rate is very high, but the amount of calculation is also so large. Currently, international standardized face expression classification includes seven classes neutral, anger, happiness, sadness, surprise, disgust and fear [5, 6].

Many researchers have proposed various methods to detect and recognize face expressions. In general face expression representation could be categorized either as holistic, analytic or hybrid methods [1][7]. In holistic approach, the whole face region is taken as data input into the face expression recognition system. Examples of holistic methods are eigenfaces, probabilistic eigenfaces, fisherfaces, support vector machines, nearest feature lines (NFL) and independent-component analysis approaches. In analytic-based approaches, local feature points on face such as the nose, the mouth, the eyes are segmented and then used as input data to the classifier while hybrid separately extracts both local and global features and combines them for recognition. However, this field still remains very challenging especially in real life applications.

Recently, Feng in 2004[8] used Local binary parameters to extract face appearance features. They used a two stage classifier, at the first stage, two expression candidates from initial seven were selected.

Author α σ ρ : School of Information Science and Engineering, Central South University, Changsha, 410083, China.
E-mail α : fedhap@yahoo.com
E-mail σ : djhuang@mail.csu.edu.cn
E-mail ρ : rimirurm@yahoo.com

At the second stage, one of the two candidate classes was verified as final expression class. In 2006, Tsai and Jan [9] used subspace model analysis to analyze the data and to recognize facial expressions. They also did some research on facial deformation problems e.g. pose or illumination variations. Nan and Youwei [10] used five classifiers and then used Dempster-Shafer (DS) classifier combination approach. They reportedly achieved a maximum accuracy of 95.7%. Wallhoff et al. in [11] discussed innovative holistic and self organizing approaches for efficient facial expression analysis. Their experiments were based on publicly available FEEDTUM database. They achieved accuracy of 61.67% by using macro motion blocks and SVM-SFFS as feature extraction and feature classification respectively. In 2008, Kotsia et al. [12] did an analysis of the effect of partial occlusion on facial expression recognition, using Gabor wavelets, Discriminant Non-negative Matrix Factorization and a shape-based method as feature extraction techniques. Whitehill et al. [13] explored an idea for recognition of facial expression in relation with intelligent tutoring system. Their idea was to automatically estimate the difficulty level of the lecture as perceived by the student as well as to determine the preferred viewing speed of the student. In 2009, Tai and Huang [14] proposed a method for facial expression recognition in video sequences. They performed noise reduction using median filter and then used a cross-correlation of optical flow and mathematical models from the facial points. Finally the features were fed to an ELMAN neural network for expression classification.

Even though many researchers have used various methods to recognize facial expressions from images and videos, utilizing multi-resolution technique to process image pixel value for recognizing the facial expressions has not been exhausted. One of the most popular multi-resolution analysis technique is the wavelet transform. Wavelet transform can be performed for every scale and translation, resulting in Continuous Wavelet Transform. Wavelet transform can also be performed at multiples of scale and translation intervals, resulting in Discrete Wavelet Transform (DWT). Since CWT provides redundant information and involves more computational effort, normally DWT is preferred [15]. In [16], Haar-like technique has been used to extract the features. Six statistical features namely variance, standard deviation, mean, power, energy and entropy were derived from the approximation coefficients of Haar-like decomposition. These statistical features were used as an input to the neural network for classifying 8 facial expressions.

Preprocessing is regarded as an important step in image processing as it helps improve the image quality by removing noise, highlighting features of interest and separating the object of interest from the background. In this work, we use morphological opening and closing operators to eliminate noise and its effects from the input image before using Haar discrete wavelet transform to extract features for the neural network classifier.

a) Proposed Method

In this paper, a face represents a set of connected regions of similar texture and intensity levels that are combined to form objects. Some of these objects are small and low in contrast while others are large and high in contrast, presenting the need to analyze them using multi-resolution processing. In order to address the curse of dimension, we map the face image model into a low level dimension system that reflects the dynamics of the human facial expression system, then use binary image processing techniques to reject noise introduced by cropping facial images through low resolution disparity calculations that consequently lead to gain of better results. Next the discrete wavelet transform is used to extract features for a neural network classifier.

The rest of the paper is organized as follows: Section 2 gives the image preprocessing, Section 3 gives detailed description on discrete wavelet transform. Section 4 gives the details of a neural network 5 presents the experiment results and analysis. Finally, we conclude in Section 6.

II. Image Pre-Processing

The first step is to acquire images from the sensor or from the database. In our experiment we use static images from JAFEE database. Image preprocessing is a significant step it transforms the image data until regions of interest, better suited for analysis are found. First we crop the facial part of the image in order to remove hair, the neck and other background details that are not central to face expression followed by histogram equalization to enhance the quality of the image. Next we use morphological operator opening and closing to eliminate noise and its effects which may have arisen from image acquisition while distorting the image as little as possible. Morphological processing compares pixels to those pixels surrounding it. They change the shape of particles by processing each pixel based on its number of neighbors and the values of those neighbors. A neighbor is a pixel whose value affects the values of nearby pixels during certain image processing functions. Morphological transformations use a 2D binary mask to define the size and effect of the neighborhood on each pixel, controlling the effect of the binary morphological functions on the shape and the boundary of a particle. Morphological opening defined as the opening of an image X by structuring element B is the erosion of X by B followed by the dilation of the result by B.

$$X \circ B = (X \ominus B) \oplus B \qquad (1)$$

Similarly closing of an image X by structuring element B is dilation followed by erosion.

$$X \bullet B = (X \oplus B) \ominus B \quad (2)$$

Erosion stage eliminates the noise from the background but increases the size of noise elements (dark spots) contained in the image because they are inner boundaries that increase in size as objects are eroded. The enlargement is countered by performing dilation on the resulting face image. Morphological opening creates some gaps within the image this is fixed by performing a closing operation on the opening. The results are given in figure1. (d) Closing has the overall effect of smoothening the image and eliminating small holes.

Figure1 (a) shows the original face, Figure1 (b) original cropped image. Figure1 (c) gives the result of opening image in (b) with a structuring element. Figure1 (d) is the result of performing a closing on results of figure1(c), it gives the net result a smoothened image after noise elimination both in the background and in the face image.

Figure 1. (a) Original test image (b) A cropped test image (c) morphological opened test image (d) morphological closing

III. Discrete Wavelet Transform

Wavelet transform has good time-frequency localization property and better energy concentration performance in addition to benefits of dimension reduction that leads to less computation complexity, multi-resolution data approximation and insensitive feature extraction. Discrete wavelet transform operates by convolving a target function with wavelet kernels, to obtain wavelet coefficients representing the contributions of wavelets in the function at different scales and orientations. DWT can be mathematically expressed by Equation 1 as:

$$DWT_{x(n)} = \begin{cases} d_j, k = \sum x(n) h_j^*(n - 2^j k) \\ a_j, k = \sum x(n) g_j^*(n - 2^j k) \end{cases} \quad (3)$$

The coefficients $d_{j,k}$ refer to the detail components in signal ***x(n)*** and correspond to the wavelet function, whereas $a_{j,k}$ refer to the approximation components in the signal. The functions ***h(n)*** and ***g(n)*** in the equation represent the coefficients of the high-pass and low-pass filters respectively, whilst parameters j and k refer to wavelet scale and translation factors.

By applying 2-D DWT to an image we decompose it into four equal sub-bands each a fourth of the original image as shown in figure 2(a). The LL, LH, HL and HH sub-bands corresponding to approximate, horizontal, vertical and diagonal matrices respectively. The LL corresponds to low frequency information from both the horizontal and vertical directions; LH corresponds to the low frequency components in the horizontal direction and high frequency in the vertical direction. HL corresponds to the high frequency components in the horizontal direction and low frequency in the vertical direction. And HH corresponds to the high frequency components in both the horizontal and the vertical direction. Taking the LL sub-band to represent the image we compress the original image into quarter its dimension, which leads to less computation complexity and recognition time. Further wavelet decomposition on the LL image generates lower dimensional multi resolution facial image. If the decomposed components continue to be decomposed, it forms a pyramid-like wavelet decomposition tree structure that might be beneficial for further analysis.

It is worth noting that the LL sub-band of an image carries the general and most important features of the face which is necessary for recognition while the high frequency carries most detailed information that happen when a person is smiling, sleeping, annoyed It and so on which are key to our face expression recognition. From figure 2(b) it is clear that the eyebrows, eyes and mouth contour in the HL are very clear. To be able to verify the impact of high frequency components on the expression recognition, we extract the low-frequency components LL and add high-frequency components HL, LH and HH. The steps are as follows:

Step 1: Two-level wavelet decomposition of training set with haar wavelet packet. Organize the low-frequency LL into column vector, denoted by X;

Step 2 : Change high-frequency components HL, LH and HH into a column vector, then respectively add them to column vector X to form feature vector X';

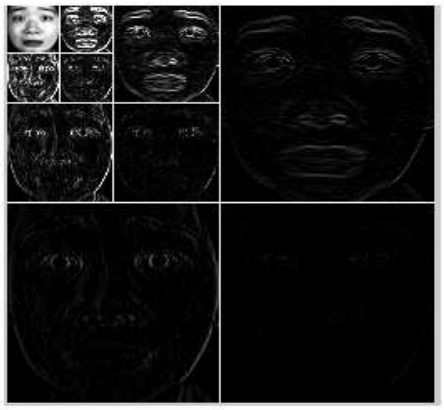
Step 3 : Feature vectors obtained from Step2 form the input to our neural network for classification.

a(n)	h(n)	h(n-1)	h(n-2)
v(n)	d(n)		
v(n-1)		d(n-1)	
v(n-2)			d(n-2)

Figure (2a) 3rd level DWT decomposition structure, n denotes the decomposition scale and a, h, v, d are approximation, horizontal, vertical and diagonal detail coefficients, respectively.

The sub image in the upper left corner is the approximation image that results from final decomposition step surrounded in a clockwise manner by the horizontal, diagonal and vertical detail coefficients that were generated during the same decomposition.

Figure 2(b) A processed image



3rd level Haar wavelet decomposition

IV. Back Propagation Neural Network (BPNN)

Back propagation method enables the network to learn a predefined set of input-output example pairs by using a two-phase propagate-adapt cycle. First an input pattern is applied as a stimulus to the first layer of network units, which is propagated through each hidden layer until an output is generated. The actual network outputs are subtracted from the desired outputs and an error signal is produced. This error signal is the basis for the back propagation step, whereby the errors are passed back through the network by computing the contribution of each hidden processing layer and deriving the corresponding adjustment needed to produce the correct output. This process repeats, layer by layer, until each node in the network receives an error signal that describes its relative contribution to the total error. Based on the error signal received, connection weights are then updated by each unit to cause the network to converge toward a state that allows all the training patterns to be encoded.

The back propagation network used consists of input layer, hidden and output layer. Before the training process the weights are initialized to small random numbers to ensure that the network is not saturated by large values of the weights and to prevent other training pathologies in addition to making sure that the network learns. The number of neurons in the hidden layer were varied between 6 and 15.In our analysis we found the system to perform well with 10 neurons. The number of neurons in the output layer equals to 6, the number of classes. The stopping criteria used the sum of squared error (SSE) (1.0) and a maximum number of epochs equals to (10000). The basic procedure for training back propagation network is embodied in the description given below [17].

Algorithm :

i. Initialize the network weights and biases
ii. Select the training pair from the training set apply input vector to the network input
iii. Sum the weighted inputs and apply activation function to compute output signal.

$$X_{pj}^{h} = \sum_{i=1}^{N} w_{ji}^{h} x_{pi} + b_{j}^{h} \Rightarrow i_{pj=} f_{j}^{h}(X_{pj}^{h}) \quad (4)$$

where w_{ji}^{h} is the weight connection from the i^{th} input unit and b_{j}^{h} is the bias. 'h' superscript refers to quantities of the hidden layer.

iv. Calculate the output of the network

$$y_{pk}^{o} = \sum_{j=1}^{L} w_{kj}^{0} i_{pj} + b_{k}^{o} \Rightarrow o_{pk} = f_{k}^{o}(y_{pk}^{o}) \quad (5)$$

where superscript 'o' refers to quantities at the output layer.

v. Calculate the error terms for the output units:

$$\delta_{pk}^{o} = (y_{pk} - o_{pk}) f_{k}^{o'}(y_{pk}^{0}) \quad (6)$$

y_{pk} is the desired output value, and o_{pk} is the actual output Followed by the error terms for the hidden units:

$$\delta_{pk}^{h} = f_{j}^{h'}(X_{pj}^{h}) \sum \delta_{pk}^{o} w_{kj}^{o} \quad (7)$$

Update weights on the output layer:

$$w_{ji}''(t+1) = w_{ji}^{h} t + \eta \delta_{pj}^{n} x_{i} \quad (8)$$

Update the weights on the hidden layer:

$$w_{ji}''(t+1) = w_{ji}^{h} t + \eta \delta_{pj}^{n} x_{i} \quad (9)$$

Repeat the above steps with all the training vectors until the error for all vectors in the training set is reduced to an acceptable value.

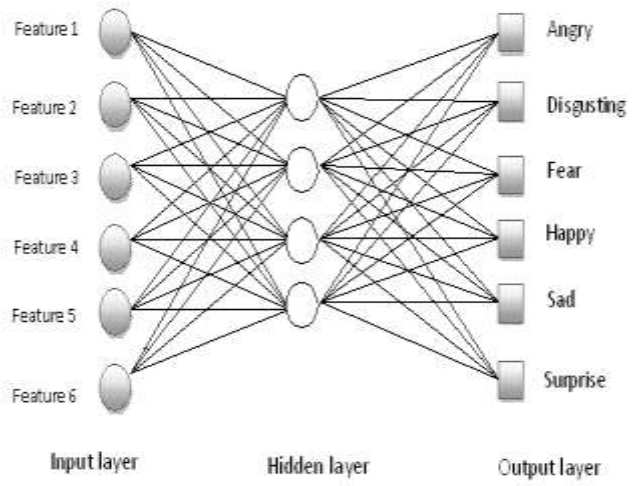


Figure 3. Illustration of the neural network training

V. Experiment Results and Analysis

To assess the validity and efficiency of our approach, the experiments are conducted on the Japanese Female Facial Expression (JAFFE) database, which contains 213 images of 7 facial expressions posed by 10 Japanese female models. Ten expressers posed 3 or 4 examples of each of six facial basic expressions (happiness, sadness, surprise, anger, disgust, and fear) and a neutral expression. Images from the JAFFE database are cropped to eliminate most of the background and some parts of the hair and chin. The sizes of the images are 256×256 in tiff format. Figure 4 shows examples of the original images in the JAFFE database. In the real-world environments, the rotation of the camera axis and the head pose variations often exist. JAFFE database presented includes these images with minor rotation of camera axis and variations in head poses. As a result, the robustness of the proposed method is evaluated.

The images depicting six different facial expressions: fear, disgusting, happiness, sadness, anger and surprise were used. The feature vectors are extracted from second level Haar discrete wavelet decomposition of the corresponding image. Beyond this level we noted that the size of images become unreasonably small and no valuable information could be extracted. The extracted image data are divided into testing and training data. In training phase two feature vectors per class were used and in the testing phase the remaining feature vector from each class was used. The images used in the testing set were not included in the training set. We use six binary NN classifiers, the data is divided into six blocks according to six expression classes, and each classifier is trained for a particular expression using one-against-all approach. The output of these binary classifiers gives the probabilities to which extent the input image belongs or does not belong to the class for which the particular classifier has been trained. After the training, we use the output generated as a feature map that provides an indication of the presence or absence of many face expression feature combinations at the input. We repeated the training procedure 30 times and got average results for the trials, with a percentage accuracy of 81%.

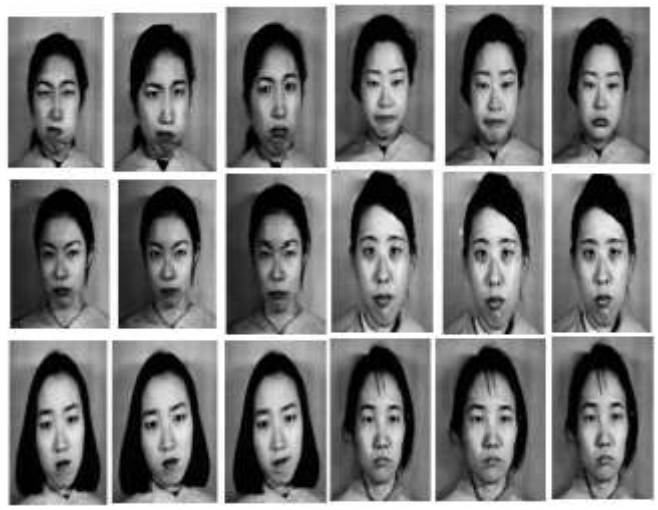


Figure 4: A sample of angry faces from JAFEE database

The results are illustrated using the confusions matrix table 1. The row corresponds to the six face expressions while the columns correspond to the

predicted emotions. The diagonals entries show the percentage of correct classification for each class and off diagonal entries show misclassification. Accuracy is computed as follows

$$\%\text{accuracy}=\frac{\text{number of correct classification classification}}{\text{number of all classification}}\times 100$$

Table 1. Confusion matrix of six basic face expressions

	angry	disgusting	fear	happy	sad	surprise	unclassified
angry	100						
disgusting		100					
fear			57				43
happy				100			
sad					27		73
surprise						100	

VI. Conclusion

In this work the high accuracy recognition system based on machine learning with a reasonable number of samples are introduced. First the input image is preprocessed, morphological operators are used to remove noise and to smoothen it the resulting binary data is used as input to discrete wavelet transform. Second level Haar wavelet is computed and the resulting feature vectors are used as input to the back propagation neural network classifier. Experiments for evaluation were carried out on JAFEE database presenting the six facial expressions, 'angry', 'disgusting', 'fear', 'happy', 'sad', 'surprise' and the results have shown that the proposed method can perform at 81% accuracy. The low results arise from unclassified data of fear class (43%)and sad class (73%). In these analysis the simplicity and robustness of the system is significant. For our future work, we plan to look into the facial expression recognition of subjects in real-time videos.

VII. Acknowledgments

We thank JAFFE Database for providing us the face images for the experiments. This work was partially supported by the National Natural Science Foundation of China (50275150) and the National Research Foundation for the Doctoral Program of Higher Education of China (20040533035, 20070533131).

References Références Referencias

1. B. Fasel and J.Luettin, "Automatic facial expression analysis": a survey, *Pattern Recognition*, 2003, 36(1): 259～ 275
2. L. Chen, H. Tao, T. Huang, T. Miyasato, and R. Nakatsu, "Recognising realistic emotions and affect in speech: State of the art and lessons learnt from the first challenge," *Proc. IEEE Workshop on Multimedia Signal Processing*, pp. 83–88, 1998.
3. L. C. De Silva and P. C. Ng, "Bimodal emotion recognition," in Proc.4th IEEE Int. Conf. on Automatic Face and Gesture Recognition,France, Mar. 2000, pp. 332–335.
4. Ekman, P. and Friesen W, "Facial action coding system: a technique for the measurement of facial movement", Palo Alto: Consulting Psychologists Press, 1978.
5. S. Dongcheng and JIANG Jieqing,"The Method of Facial Expression Recognition Based on DWT-PCA/LDA", 2010 3rd International Congress on Image and Signal Processing.
6. WANG Zhiliang, LIU Fang,WANG Li. "Survey of Facial Expression Recognition Based on Computer Vision[J]", Computer Engineering, 2006, 32(11),231～33.
7. Elham Bagherian, Rahmita Wirza O.K. Rahmat, *"Facial feature extraction for face recognition: a review"*,*Information Technology*, *2008*. *ITSim 2008*. *International Symposium* on 09/*2008*; 2: *1-9*.
8. Xiaoyi Feng, "Facial expression recognition based on local binary patterns and coarse-to-fine classification", in Proceedings of the The Fourth International Conference on Computer and Information Technology (CIT'04). Pages: 178 – 183. Year: 2004.
9. Tsai, P.H. Jan, T, "Expression-invariant face recognition system using subspace model analysis", in IEEE International Conference on Systems, Man and Cybernetics, 2005.
10. Zhang Nan, Zhang Youwei, "Inducement analysis in facial expression recognition". In The 8th International Conference on Signal Processing 2006, Volume III. Year: 2006.
11. F. Wallhoff, B. Schuller, M. Hawellek and G. Rigoll, "Efficient recognition of authentic dynamic facial expressions on FEEDTUM database". In IEEE Conference on Multimedia and Expo (ICME'06), Pages: 493-496, Year: 2006.

12. Irene Kotsia, Ioan Buciu, Ioannis Pitas, “An analysis of facial expression recognition under partial facial image occlusion”, In transanctions on Image and Vision Computing, Volume 26, Issue 7, Pages: 1052-1067, Year 2008.
13. Whitehill, J.; Bartlett, M.; Movellan, J., “Automatic facial expression recognition for intelligent tutoring systems”, In proceedings of IEEE computer scociety workshop on Computer Vision and Pattern Recognition, Pages: 1-6, Year 2008.
14. Shenchuan Tai, Hungfu Huang, “Facial expression recognition in video sequences”, In Proceedings of the 6th International Symposium on Neural Networks: Advances in Neural Networks - Part III, Pages 1026- 1033, Year 2009.
15. Murugappan, M., Rizon, M., Nagarajan, R., and Yaacob, S., FCM clustering of human emotions using wavelet based features from EEG. Trans. Biomed. Soft Comput. Hum. Sci. IJBSCHS 14(2):35–40, 2009.
16. Satiyan, M., and Nagarajan, R., Recognition of facial expression using Haar-like feature extraction method. Proceedings of the 3rd IEEE International Conference on Intelligent and Advanced Systems (ICIAS), Kuala Lumpur, Malaysia, pages 1-4, Year 2010.
17. James A. Freeman. David M. Sukapra." Neural Networks Algorithms, Applications and Programming Techniques", Addison-Wesley Publishing Company, New Yolk.1991.

Global Journal of Computer Science and Technology Volume XII Issue VII Version I April 2012
14

This page is intentionally left blank

Global Journal of Computer Science and Technology
Volume 12 Issue 7 Version 1.0 April 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Search-Ability a Domain Quality Factor for Web Software Applications

By B. Narayana Babu & C.S. Ramanathan
Sri Sankara Arts and Science College, Kanchipuram, India

Abstract - As organizations become aware of the strategic importance of e-commerce they will also become aware of the need of quality Web sites. In early years the World Wide Web was originally designed to present information to Web surfers using simple sites that consists of hyper linked text. But, Modern Web applications run large-scale software applications for e-commerce, information distribution, entertainment, and numerous other activities. The factors that constitute software quality in traditional data processing are well defined. However, it is necessary to have a full understanding about the quality in the context of World Wide Web. This paper identifies a new quality factor, searchability for the World Wide Web with the checklist of enablers. This factor enables the Web site developers and evaluators to create quality Web sites.

Keywords : *World Wide Web, Internet, quality factor, searchability, checklist of enablers.*

GJCST Classification: *H.3.5*

Strictly as per the compliance and regulations of:

Search-Ability a Domain Quality Factor for Web Software Applications

B. Narayana Babu [α] & C.S. Ramanathan [σ]

Abstract - As organizations become aware of the strategic importance of e-commerce they will also become aware of the need of quality Web sites. In early years the World Wide Web was originally designed to present information to Web surfers using simple sites that consists of hyper linked text. But, Modern Web applications run large-scale software applications for e-commerce, information distribution, entertainment, and numerous other activities. The factors that constitute software quality in traditional data processing are well defined. However, it is necessary to have a full understanding about the quality in the context of World Wide Web. This paper identifies a new quality factor, *searchability* for the World Wide Web with the checklist of enablers. This factor enables the Web site developers and evaluators to create quality Web sites.

Keywords : World Wide Web, Internet, quality factor, searchability, checklist of enablers.

I. Introduction

The World Wide Web was originally designed to present information to Web surfers using simple sites that consist of hyper linked text documents. According to Bevan [BEV98] "Web sites provide a unique opportunity for inexperienced information providers to create a new generation of difficult to use systems". Modern Web applications consists of diverse components including traditional and non traditional software's, interpreted scripting languages, plain HTML files, mixtures of HTML and programs, databases, graphical images, sounds and complex user interfaces. Most of these Web sites are developed by enthusiastic beginners. They have the perception that, a *quality* site is one that demonstrates the latest multimedia and animation effects. As the result the site has been difficult to read, find, navigate and takes long time to load.

As such, engineering an effective Website requires large teams of people with very diverse skills and backgrounds. These teams include programmers, graphic designers, usability engineers, information layout specialists, data communications and network experts, and database administrators [ELI02].

Several factors inherent to Web development contribute to the quality problem. Developers build Web-based software systems by integrating numerous diverse components from disparate sources, including custom-built, special-purpose applications, customized off-the-shelf software components, and third-party products. In such an environment, systems designers choose from the potentially numerous components and they need information about the various components' suitability to make informed decisions about the software's required quality attributes.

Author α σ : Assistance Professor, Department of Computer science Sri Sankara Arts and Science College, Kanchipuram, India.
E-mail α : nbslnc@yahoo.com
E-mail σ : tpt_auctc@yahoo.co.in

II. Searchability – A Domain Quality Factor for Web Applications

As new domains evolve and are understood there is a need to review our interpretation of quality in those new domains [RON00] [JEF02].

a) Searchability

The majority of traffic to meet Web sites comes through the major search engines and directories. The NUA analysis 2010, states the importance of search engines by noting that 55% of Internet shoppers made purchase at sites they found through the major search engines and only 9% who bought at sites found through banner ads [ROB02]. So, our site must be identifiable by the search engines and visitors. In order to do that we must submit our page using a handy submission service or a software program that will submit pages to more search engines.

b) Search Engines

The search engines allow the users to find pages on our site related to a specific keyword or phrase. When we submitting a Web site to the search engines the search engine look at each page of a Web site separately from the other pages, each page stands on its own. Which means that each and individual page of your Web site has its own value according to the Search engines. So, each individual page needs to be optimized separately when working with search engines. To optimize means to create tags and texts in accordance with a particular search engine's unique likes and dislikes, in an attempt to get your site placed higher in the rankings of that engine. So, your tags, such as <TITLE>, <META description>, <ALT>, and so forth, will be designed for that one particular page whereas another page will have different tags and a different focus.

The search engines use *spiders* to index the Web pages. Spiders are sophisticated software

programs that crawl the Web. Index the sites, and build and index of Web pages that we access when we search. Web site owners submit their URL's to the engine, and a response, the engine send its spider to the pages to index them [ROB02].

When search engines crawl and spider Web pages, they also follow links from those Web pages, so they often index those pages as well. Therefore. You can see the importance of providing links to all your important Web pages on your main page. In our analysis we discuss the following search engines, with links to find the engines and strategies for better ranking.

- Google
- AltaVista
- Excite
- HotBot
- Lycos

The following table shows the ranking of web sites for the phrase "image compression".

Table 1. Top 10 Ranking of Web Sites in Each Search Engine for the Phrase "Image Compression"

Web Site	AltaVista	Excite	Google	Hotbot	Lycos
www.xat.com	1	1		1	
www.cs.dartmouth.edu		2	5	3	
www.bitjazz.com		3			3
www.faqs.org	5	4	1 & 2		1
www.cvisiontech.com		5			
www.c3.lanl.gov	2		3 & 4		
www.iterated.com					4
www.spinwave.com	7			2	
www.jeffdavis.net	3				
www.acm.org	4				
www.dip.ee.uct.ac.za				5	
www.debugmode.com					2
www.ricazip.com					5
www.Webreference.com				4	

Though, each search engine has it's own strategy and rules for ranking. We found the following checklists of enablers are common to all search engines.

Checklist of Enablers:

- ✓ The single important thing we can do in terms of search engine strategy is to choose the best keyword. Choose different keywords depending on the content of the page.
- ✓ Use keyword phrases rather than single keywords.
- ✓ Use all variations of your keyword phrase in the tags and text on your page. Because, some engines are case sensitive.
- ✓ Specify the keyword in bold in the body of the text.
- ✓ When deciding where to place your important keyword in your site, you also need to consider the page keyword weight. It refers to the number of keywords that appear on your page in relation to the total number of words in the page. A good general keyword weight is 3 to 5 percent of viewable text.
- ✓ Use the keyword phrase in the beginning of the tag.
- ✓ The most important tag on your page is <TITLE> tag. So, be sure to use the keyword phrase in the beginning of the tag.
- ✓ Move the <TITLE> tag to the beginning of the page.
- ✓ The <TITLE> tag's optimal number of characters is around 75 characters.
- ✓ The description tags optimal number of characters is around 128 characters.
- ✓ Use the keyword phrase in the Headline tags.
- ✓ Don't use ALL CAPS in the tags.
- ✓ Use <META> tag and specify the keywords and description attributes with the keyword phrase in it.
- ✓ Avoid the use of irrelevant <META> tags.
- ✓ Avoid the use of <META> refresh tag
- ✓ Specify the link with the related keyword of the page.
- ✓ Purchase Domain name for your Web site with relevant Keyword phrase on it.
- ✓ Name your page with the keyword phrase on it.
- ✓ Name image files and multi-media files with the keyword phrase on it.
- ✓ Use ALT attribute for images and specify the keyword phrase.
- ✓ Always use <NOFRAMES> version for the page containing <FRAMES>
- ✓ Use contact information in your Web site.
- ✓ Place a copyright notice on your page.
- ✓ Minimize the use of JavaScript code. Because, it can push your keyword containing text further down on the page.

- ✓ If you want to use the Java Script, move the JavaScript to a separate external .js file, and then reference that file in the <HEAD> section of your page.
- ✓ <SCRIPT LANGUAGE = "JavaScript" SRC = "namefile.js">
- ✓ </SCRIPT>.
- ✓ If you decide to use dynamic pages (i.e., the URL contains a '?' character), be aware that not every search engine spider crawls dynamic pages as well as static pages. It helps to keep the parameters short and the number of them small.
- ✓ Don't design pages that take more than 60 seconds to load.
- ✓ Minimize the use of image maps. Because it will push down your keyword containing text further down to the page.
- ✓ If you are using image maps, provide corresponding text link. It will help the search engine to find the text.
- ✓ Always link relevant sites.
- ✓ Add more content if the top ranking sites for that keyword phrase contain more than this page. Reduce the content if the competitor uses fewer words.

Characteristics for Searchability :

The following two characteristics support the *searchability* of the web site.

Trace-ability (Traceability) : It supports potential visitors by enabling them to find a Web site.

Accessibility : It supports the easy retrieval and the ease of access of the Web site.

III. Conclusion

This paper explains the manner in which Web sites are currently developed without reference to *quality* considerations. Also analyses the quality considerations for the Web sites and introduces new quality factor *searchability* with the checklist of enablers. This domain specific quality factor when combined with the core software quality factors can be used as an essential component for a quality accreditation system for Web sites. This paper also shows that as new domains evolve and are understood there is a need to review our interpretation of quality in those new domains.

References Références Referencias

1. [BEV98] Bevan. N and Macloed, "Usability Issues in Web Site Design", Proceeding's of UPA'98, Washington, 1998.
2. [ELI02] Elizabeth Hendricken and Martin Fowler, "The Software Engineering of Internet Software", IEEE Software, March/April 2002. pp 23,24.
3. [JEF02] Jeff Offut, "Quality Attributes of Web Software Applications", IEEE Software, March/April 2002. pp 25-32.
4. [ROB02] Robin Nobles and Kerri-Leigh Grady, "Web Site Analysis and Reporting", Prentice Hall India, New Delhi, 2002. pp 285,311,357,415,441.
5. [ROB03] Robert Hall, "Support for Web Service Developers", ORACLE Magazine, March/April 2003, pp 63,64.
6. [RON00] Ronan Fitzpatrick, "Interpreting Quality Factors for World Wide Web",

This page is intentionally left blank

Global Journal of Computer Science and Technology
Volume 12 Issue 7 Version 1.0 April 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Analyzing & Identifying Saas for Development of a Project by Calculating its Reputation

By Dr.Bora Rama Rao
Dr Iankapalli Bullayya college of Engineering

Abstract - Assessing the quality of external software before integrating it in to the project development is very challenging now days. As IT industry is moving towards newly evolving tool named SaaS(Software as a Service) , the risk of integrating the external software to the project development has been increased . Presently integration of external software is going on, but they use the trad itional way of collecting the feedbacks to identify whether to use that external software into the project or not, which may produce an unfair results at the end of project deployment. So in this perspective we are going to propose an automated framework to rate and select a service by identifying quality and reputation .And we mainly focused on addressing the risk in proposing external software by using quality and reputation of it.

Keywords : *Software as a Service, Quality, Project development, Automation, risks.*

GJCST Classification: *D.2.8*

Strictly as per the compliance and regulations of:

Analyzing & Identifying Saas for Development of a Project by Calculating its Reputation

Dr.Bora Rama Rao

Abstract - Assessing the quality of external software before integrating it in to the project development is very challenging now days. As IT industry is moving towards newly evolving tool named SaaS(Software as a Service) , the risk of integrating the external software to the project development has been increased . Presently integration of external software is going on, but they use the trad itional way of collecting the feedbacks to identify whether to use that external software into the project or not, which may produce an unfair results at the end of project deployment. So in this perspective we are going to propose an automated framework to rate and select a service by identifying quality and reputation .And we mainly focused on addressing the risk in proposing external software by using quality and reputation of it.

Keywords : Software as a Service, Quality, Project development, Automation, risks.

I. Introduction

Software Engineering can be stated in Industrial terms as the application of a quantifiable, systematic, disciplined approach to the development, process, and maintenance of software. As software industry has huge competition it has shaped a strong motivation for developing solutions to support more responsive and more competitive businesses. Even with long-standing success of COTS (commercial off-the-shelf) software as a time- effective alternative to custom "in-house" developed solutions is still being compromised by the implicated cost of ownership, installation and maintenance time, and effort. That's the reason why software industry has started moving toward a new kind of software delivery model called SaaS(Software as a Service) and which made the things easy to install, maintenance-free, and money- spinning. In Software as a Service (SaaS) software delivery model the software is delivered on-demand and priced on-use, which made it to be widespread implementation of fast Internet access, combined with the widespread acceptance of SOA based solutions.

Fig 1: The Growth of SaaS

SaaS has gained popularity by reducing the cost of tenure and alleviating the burden of software installation and maintenance. SaaS contributions has expanded dramatically as some of the enterprises have started to outsource their software infrastructure and development projects to SaaS vendors, and the competition has been increased even among vendors of traditional on premises software as in fig 1.

In the world of Software development using service delivery by SaaS model the quality of the software and software provider's credibility is tough and risky. So, the addition of external software in project development is challenging. In this paper risk management has been addressed in situation of project development using external software overhaul components. Reputation must be computed on the basis of fair and objective feedbacks. Most of the works that addressed until now are on evaluating the fairness of existing Feedbacks. Work in this paper focuses instead on the process of generating objective and fair feedbacks. Feedback can be individual since it is based on consumers' "personal" expectations and opinions. Consumers may have an obstructed view of a service reputation systems are prone to attacks by malicious consumers who may give false ratings and subvert service reputation. Consumers may have little incentive to leave a feedback. In this perspective a framework an automated eminence and Reputation based scaffold for service rating and selection has been proposed as in Fig.2.

Author α : Associate. Professor, Engineering Physics Dept, Dr. lankapalli Bullayya college of Engineering (for women), Visakhapatnam, A.P, INDIA.
E-mail : ram65mt@yahoo.com, balasriram1982@gmail.com

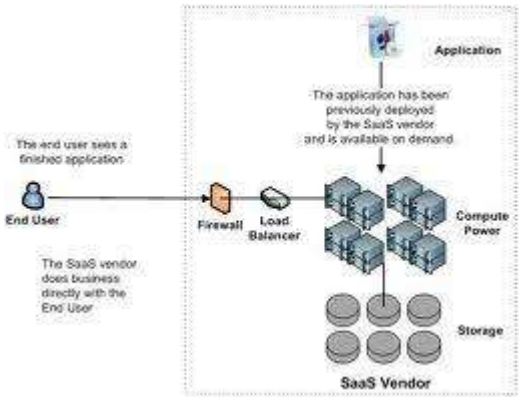


Fig.2 : The Role of Software Vendors in Implementing SaaS.

The challenges of this paper are:

a) In order for a reputation mechanism to be fair and objective, it is essential to compute reputation on the basis of fair and objective feedbacks.
b) The simulation results have demonstrated that the devised system has successfully met our primary objectives and can be an important component in a risk management strategy for software development with SaaS.
c) A computational model is provided to objectively evaluate the delivered service based on the actual measurement of the conformance of the execution quality to the contracted SLA. A novel algorithm is also devised to automate the rating process based on the expectancy-disconfirmation theory from market science.

II. Related Work

What is the main correlation stuck between "reputation" and "trust"? The major difference between reputation and trust can be illustrated by the following statements: (a) "Because of your good reputation I trust you" (b) "I trust you despite your bad reputation." Here the reputation is a collective measure of trustworthiness and is measured based on the referrals or ratings from other members in a community. According to A .josang and R. is mail, reputation is believed about a person's or thing's character or standing. Hence, trust for an individual is measured from the personal reputation and combination of received referrals, as in the Fig 3.

In a centralized reputation management system, the synthetic rating of QoS of web services is aggregated by each rating in the community. To avoid the inapt evaluation by dishonest consumers, it need identify the reputable and disreputable members with their historical comments.

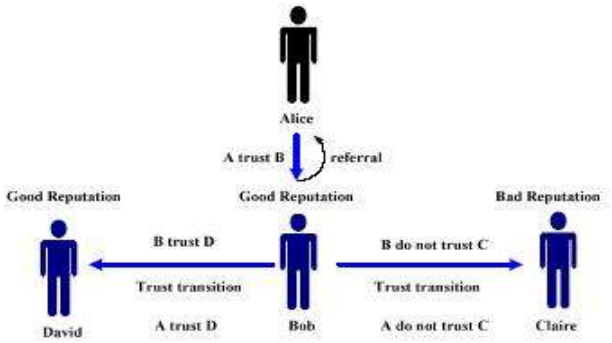


Fig 3 : A Transitive model for consumer reputation

Our idea is that consumer reputation is decided by the historical quality of comment, that is, more positive comments gain higher reputation, versa. In other words, lower reputations will worse his/her performance rating on QoS evaluation of web services. When consumers jointing the voting activity can raise their reputation by positive comments and avoid the negative comments. In this work, we proposed a centralized reputation measure for quantifying consumer reputation to properly select the service alternatives, as illustrated in Fig 4.

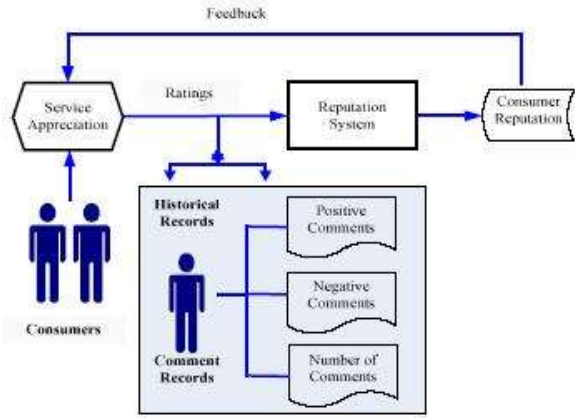


Fig 4 : Consumer Reputation Measure

III. Architecture of the Proposed

For selection of the service many previous works have measured the reputation and quality of the software, but the measurement has been done using some manual tools but none have considered the service rating process in the form of automation. WE introduce a framework for selecting and rating software to provide software service. The important point of the framework which is proposed is to automate both the rating and selection software services which is potentially increasing the objectivity of the service quality reports and concentrating on time- consumption and which finally reduces the risk associated utilization of external software services in development projects.

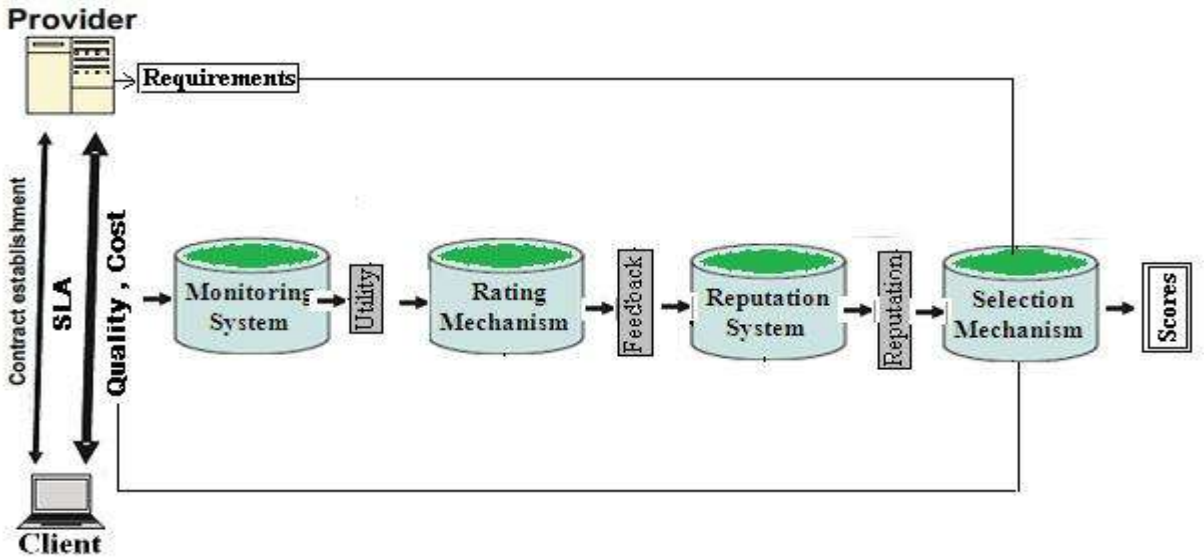


Fig 5 : System Architecture of the frameworks

While determining a service's suitability to a particular user's preferences in terms of quality and cost the service selection algorithm acts as a user-centric and reputation-aware service recommender. In order for a reputation mechanism to be fair and objective, it is essential to compute reputation on the basis of fair and objective feedbacks. Our work focuses instead on the process of generating objective and fair feedbacks, while most of the works that addressed this latter issue are on evaluating the fairness of existing feedbacks. Here concentrated the calculation of the reputation on works in the area of Service Level Agreement (SLA) monitoring where a computational model is provided to neutrally assess the delivered service based on the actual measurement of the execution quality to the contracted SLA.

In this paper we proposed a framework which has four major modules like Consumer, SLA (Service Level Agreement), Service Providers and Reputational System. Consumer can start the selection based on the trustworthiness features. Consumer selection information will be stored inside database like reputation table. SLA maintains some of the requirements about that particular service. These requirements can be co-inside with SLA requirements and for those services only the service certificate will be approved and that Certificate can be used as Trustworthiness certificate. The services which are provided by SLA can also be present in the service providers itself. User can be satisfied with certified services or trustworthy services. All the user behaviors features can be located inside the trustworthy services. To start the selection at the consumer side we should place the all the features inside that particular service. Reputation can be defined based on the frequent item selection procedure to define the utility measurement identification. Based on utility measure the feedback about that particular service will be defined. The proposed reputational framework is as shown in fig 4.

And the functional requirements of the proposed frame work will be as Enter Consumer Details, Update Consumer Required Services, and Enter Service Provider Details, Service updated to SLA, Retrieve Services, Select Service, Utility Measure of Service, Rating Function, Retrieve Feedback, Consumer Preference Updated, Select service and Calculate Score. An empirical study of the risk factors related to the development using external software (COTS-like) components along with associated risk reduction activities has been reported in. It showed that risk reduction at software selection time is negatively correlated with occurrences of most project development- related risks. In fact, selection must be driven by quality constraints, with selection time evaluation of component quality and choice of appropriate service providers all essential to successful integration. However, in practice, the evaluation of service quality cannot be performed until the service is acquired. Consequently, quality evaluation is typically limited to the evaluation of quality offers by comparing the quality level that providers promise to the quality requirements. Compliance cannot be guaranteed at selection time, so it is essential to choose a provider that is trusted to respect its commitments

IV. Design & Implementation of the System

a) Designing of the Framework

In the situation of software, design is problem solving process whose objective is to find and describe a way to find and depict the way to implement the functional requirements while respecting the constraints imposed by the non functional requirements and by

adhere to general philosophy of good quality. The goal of the design process is to produce a model or representation of a system which can be used later to build that system and use this model to build the overall system.

Design is concerned with identifying software components specifying relationships among components. Specifying software structure and providing navy print for the document phase. Modularity is one of the advantageous properties of large systems. It implies that the system is divided into several parts. In such a manner, the interaction between parts is minimal clearly specified. Design will explain software components in detail. This will help the implementation of the system. Moreover, this will guide the further changes in the system to satisfy the future requirements.

Class Diagram contains the following elements:

- A class which represents entities with common characteristics or features attributes operations and associations.
- Association, which represent relationship between two or more classes where relationships have common characteristics or features like attributes and operations as in Fig.6.

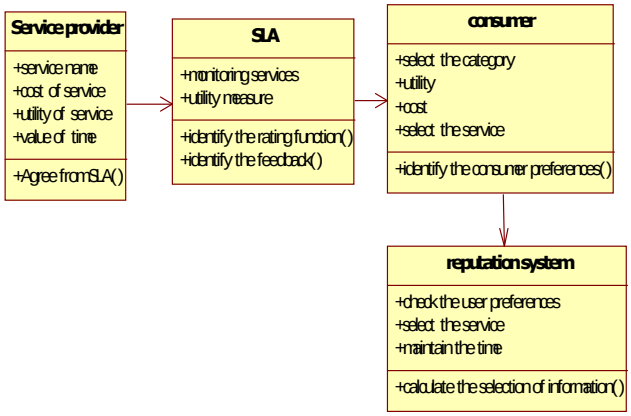


Fig 6: Inter-operational Class diagram for framework

Sequence Diagrams represents the interactions between classes to achieve a result such as a use case. The sequence diagram lists objects horizontally and time vertically, and models these messages in excess of time. In this paper the designing of the sequence has been done on both service provider and the consumer and they were described as in the Fig.7 and Fig .8

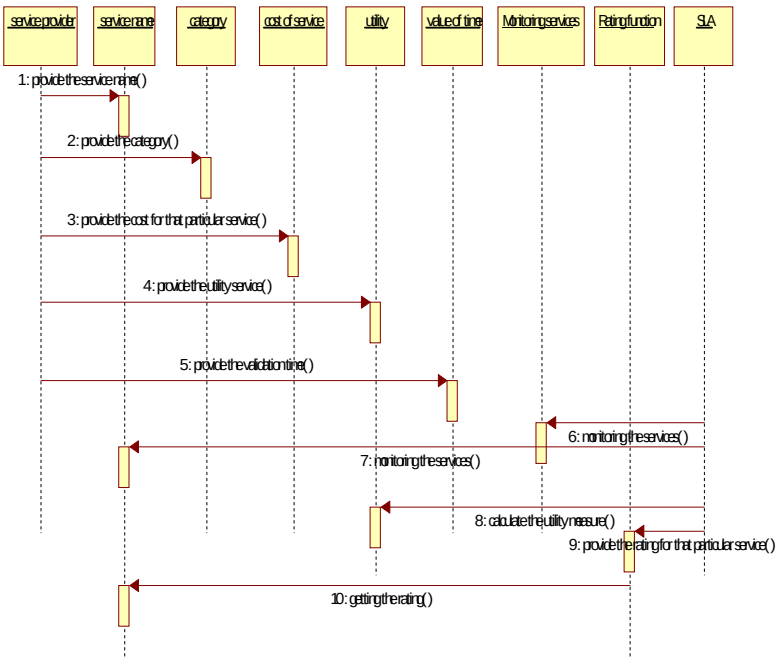


Fig 7: Inter-operational Sequence diagram for the Service Provider

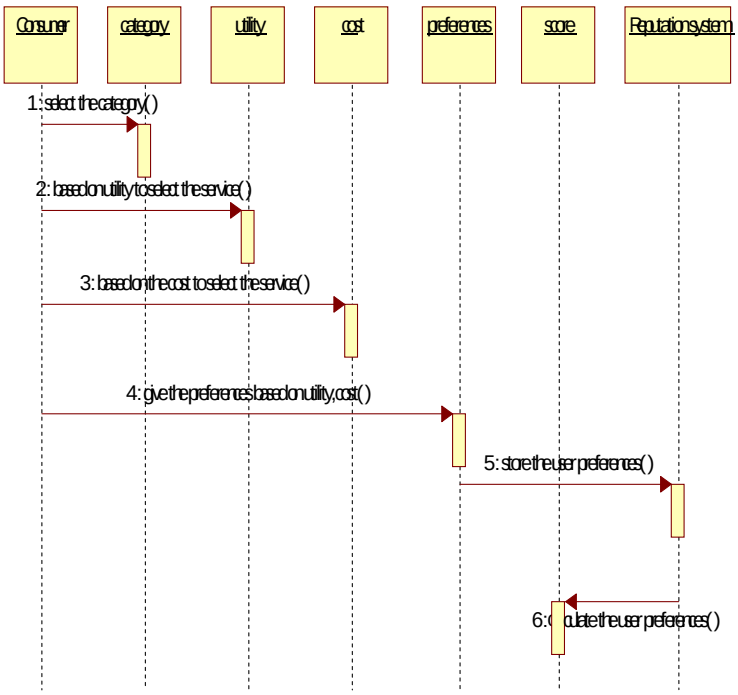


Fig 8 : Inter-operational Sequence diagram for the Consumer

b) Flowchart Representation of Framework

The Flowchart representation of the framework is detailed below:

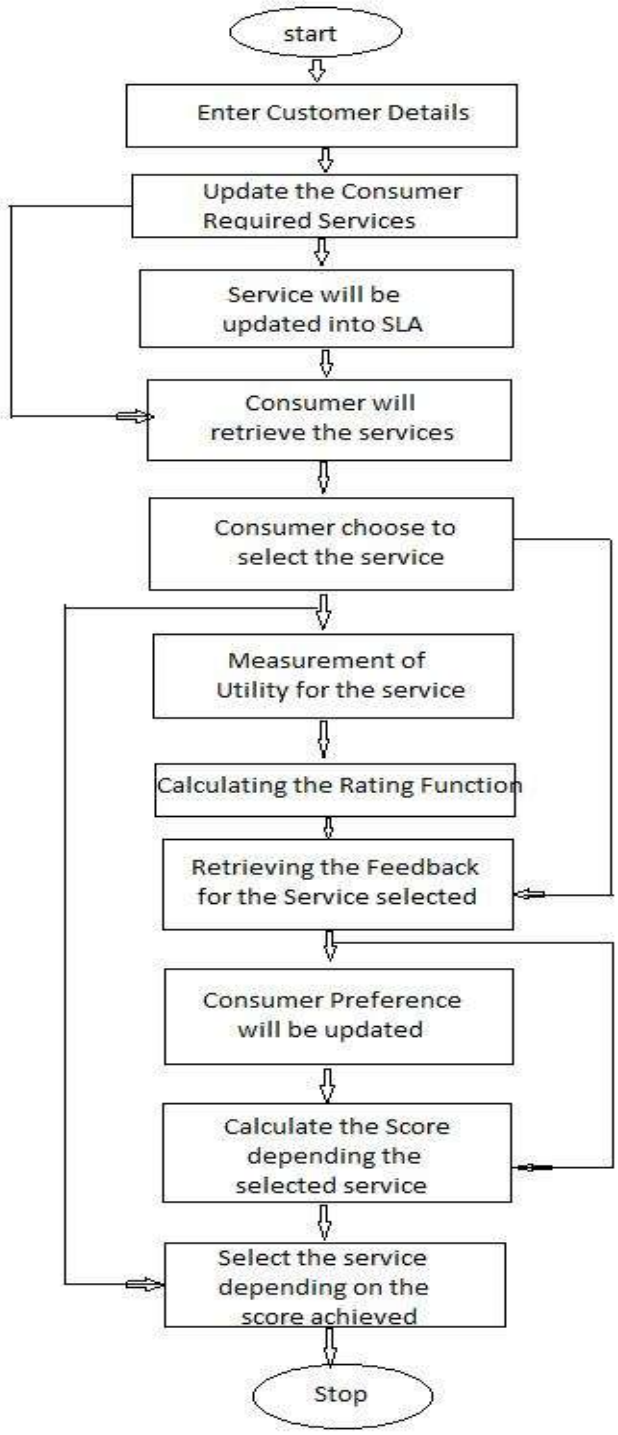


Fig.9 : Flowchart Representation of the System

V. Results

Obtained results of the system are displayed as screen shots in the following section.

Fig.10 : Selection of required service by consumer

Fig.11 : Updating the type of service provided by service provider

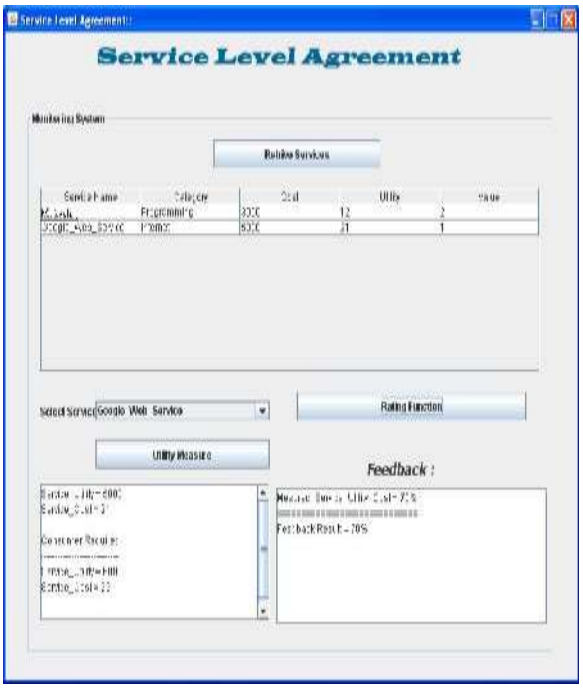


Fig.12 : The SLA between Consumer and Service Provider

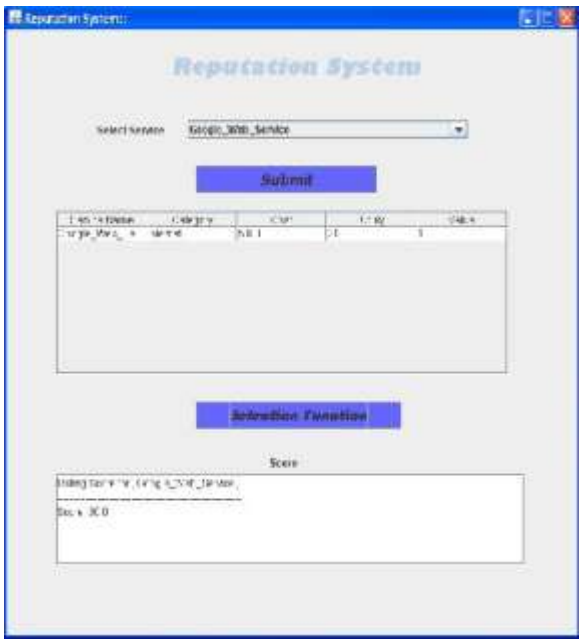


Fig.13 : Calculating the Score of the service using Reputation System

VI. Conclusion

In this paper we addressed the risk to incorporate third party software for project development. As we identified the integration of it is very risky we proposed an outstanding automated framework to rate and select a service by identifying quality and reputation. We highlighted the framework by adding enhanced features like consumer, SLA, Service Provider and Reputation System which made as added additional advantage in rating and selecting the software to be used for integration. The proposed framework have accomplished in confining the service behaviors and translating them into probable customers choice.

References Références Referencias

1. Jøsang, R. Ismail, C. Boyd, "A survey of trust and reputation systems for online service provision", Decision Support Systems, Vol.43, No.2, 2007, pp.618-644.
2. J. Li, R. Conradi, O.P. Slyngstad, M. Torchiano, M. Morisio, and C.Bunse, "A State-of-the Practice Survey of Risk Management in Development with Off-the-Shelf Software Components," IEEE Trans.
3. T. SaaS, "Trust Saas: Putting the Trust in Software as a Service (SaaS)," http://trustsaas.com/, 2008.
4. J. Skene, F. Raimondi, and W. Emmerich, "Service-Level Agreements for Electronic Services," IEEE Trans. Software Eng., vol. 36,no. 2, pp. 288-304, Mar./Apr. 2010, http://doi.ieeecomputersociety.org/10.1109/TSE.2009.55.
5. A.Keller and H. Ludwig, "The WSLA Framework: Specifying and Monitoring Service Level Agreements for Web Services,"J. Network and System Management, vol.11, no. 1, pp. 57-81, 2003.
6. Software Eng., vol. 34, no. 2, pp. 271-286, Mar./Apr. 2008.
7. J. Anselmi, D. Ardagna, and P. Cremonesi, "A QoS-Based Selection Approach of Autonomic Grid Services," Proc. Workshop Service-Oriente Computing Performance: Aspects, Issues, and Approaches, 2007.
8. L. Zeng, B. Benatallah, A.H. Ngu, M. Dumas, J. Kalagnanam, and H. Chang, "QoS-Aware Middleware for Web Services Composition," IEEE Trans. Software Eng., vol. 30, no. 5, pp. 311-327, May 2004.
9. L.-H. Vu, M. Hauswirth, and K. Aberer, "QoS-Based Service Selection and Ranking with Trust and Reputation Management," Proc. 13th Conf. Cooperative Information Systems, 2005.
10. J. Skene, F. Raimondi, and W. Emmerich, "Service-Level Agreementsfor Electronic Services," IEEE Trans. Software Eng., vol. 36,no. 2, pp. 288-304, Mar./Apr. 2010, http://doi.ieeecomputersociety.org/10.1109/TSE.2009.55.
11. J.R. Douceur, "The Sybil Attack," Proc. First Int'l Workshop Peer-to-Peer Systems, 2002.
12. J. Skene, A. Skene, J. Crampton, and W. Emmerich, "The Monitorability of Service-Level Agreements for Application-Service Provision," Proc. Sixth Int'l Workshop Software and Performanc, pp. 3-14, 2007.
13. BelGOnet,"Service Level Agreement,"http: //www.belgonet.com/website/UK/Service_Level_Agreement_UK.pdf, 2008.

Global Journal of Computer Science and Technology
Volume 12 Issue 7 Version 1.0 April 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Heterogeneous Tree Based Authenticated Group Key Transfer Protocol

By A.B.Surekha & C.Shoba Bindu
Jawaharlal Nehru Technological University,Ananatapur

Abstract - Message passing from one source to another has become a key for many upcoming technologies. This is already achieved by introduction of topics of KEYS, AUTHENTICATIONS etc. Secret key transfer is being done presently by using mutually trusted key generation centre (KGS). By this selection of session key by which encryption is done for information passing is selected. This paper discusses about the advancement of this technology by extending this service to group instead of a single key. The whole group with authenticated users can access the information. The proposed protocol considers the heterogeneity of the peer resources as QOS factor in key generation phase and shared key mechanism as primary process to achieve security in group key sharing.

Keywords : *GKMP, GKTP, P2P, Group key, QoS, Security.*

GJCST Classification: *C.2.1*

Strictly as per the compliance and regulations of:

Heterogeneous Tree Based Authenticated Group Key Transfer Protocol

A.B.Surekha [α] & C.Shoba Bindu [σ]

Abstract - Message passing from one source to another has become a key for many upcoming technologies. This is already achieved by introduction of topics of KEYS, AUTHENTICATIONS etc. Secret key transfer is being done presently by using mutually trusted key generation centre (KGS). By this selection of session key by which encryption is done for information passing is selected. This paper discusses about the advancement of this technology by extending this service to group instead of a single key. The whole group with authenticated users can access the information. The proposed protocol considers the heterogeneity of the peer resources as QOS factor in key generation phase and shared key mechanism as primary process to achieve security in group key sharing.

Keywords : GKMP, GKTP, P2P, Group key, QoS, Security.

I. Introduction

Every message under transformation ought to have security provided to it. So, for providing high security, we consider 2 issues namely (1) *Message Confidentiality*: Only the authenticated and intended user should read the message and (2) *Message Authentication*: The receiver should be assured that the sent message is from authenticated sender and the message is not altered in the middle.

Here the work of KGS starts. It should provide a one-time session key to achieve the above 2 issues of key exchange. So, KGS distributes the secret key to all intended users with confidentiality and authentication. We can see from [5] the 2 types of key establishment protocols namely Key transfer protocols, Key agreement protocols.

Apart from this the KGS helps in selecting the secret key and transport them to all communication entities secretly. These session keys are determined by all communication entities where the most commonly used is Diffie-Hellman (DH) key agreement protocol [12].

Public keys of the communication entities play a key role in this protocol. They are exchanged to fix the value of session key. As the public key itself does not provide authentication, uses a digital signature. But the only drawback is that this is on whole applicable only two 2 users but not to a group. The importance of group key is found here as everyone ought to have it. This group key management protocol can be of 2 categories. Centralize group key management protocols, where the whole group is managed by a Group Key generation. Distributed group key management, where each individual manages the generation of key rather than a group key distribution. Of the both key management protocols, we use Centralized group key management the most. It was proposed by Harney et al[15] which takes O(n) where n indicates the size of group participating in the generation of key id. In addition to this, to update this group key either adding or editing the users, we have hierarchical structure based group key protocols [10],[22],[27].

Author α : M.Tech.,Department of CSE, JNTUACE, Anantapur,INDIA. Email : rekha.pyngas77@gmail.com

Author σ : Associate Professor, Department of CSE, JNTUACE, Anantapur,INDIA. Email : shobabindhu@gmail.com

II. Related Work

We have Fiat and Naor[14] introducing a k-resistant protocol. Using this security to about k users is provided with O(k log k log n) keys and server broadcasting $O(k^2 \log^2 k \log n)$ messages per rekeying. EBS (Exclusion Basis System) proposed by Eltoweissy et al.[13] is a combinatorial formulation which helps users to switch between number of keys needed to be stored and number of messages to be transmitted. All this is for key updating so that solution to collusion is provided.

In the previous days, this group generation management protocols involved the naturally generalized DH key agreement protocol. Many examples can be quoted like Ingemarsson et al. [18], Steer et al. [28], Burmester and Desmedt [9],and Steiner et al. [29] . Later, in 1990s, Steiner et al[29] came forward with extension of DH naming it as DH key exchange[29] and in 2001, name was changed to authentication services[6].

Later from 2006, there was a drastic advancement in this group key generations. In the very year of 2006, Bohli[8] proposed a framework for group key generation agreement which is intended to provide security opposing harming participators and active unauthenticated users at every point in the network. In 2007, Katz and Yung [19] proposed the first constant-round and fully scalable group DH protocol which is provably secure in the standard model. Above all, the key feature of group DH is to generate a secret group key by a standardised group like KGS other than relying on members inside.

The next advancement in providing security is identifying the intruders present inside the network. For that, Tzeng [31] provided a conference key agreement protocol with the assistance of discrete logarithm (DL). Each user in the group requires having nm power polynomials with n representing number of participants. Later,. in 2008, Cheng and Lain [11] modified Tseng's conference key agreement protocol based on bilinear pairing. In2009, Huang et al. [16] proposed a no interactive protocol based on DL assumption to improve the efficiency of Tseng's protocol.

All the proposals made and developed till now are good. But one main problem is the time constraint. Since this key agreement involves all the communication entities, takes a lot of time for decision. So to reduce this, we have 2 different solutions. (1)All the communication entities assuming that there is an offline server active all the time and decides the secret key with this assumption.[4],[14],[25,][3]. (2) All the communication entities assuming that an online server is in active state.

Of the two, the 1st one is called key redistribution scheme. In this schema, offline users are provided with a secret piece of information created by a trusted group .But the backhand of this approach is that every server has to store a lot of secret keys and information. So we came to the 2nd approach [20] . It's working is almost similar to IEEE 802.11i standard [17] . Here, an online server votes for a group key and transmits to every group member.

Even though they employ same methodology, there is a slight difference. Instead of encrypt in the group temporal key(GTK) by Key encryption key(KEK) and individually saying the secret key information to each user, here in this approach, the information of group key is also said to all user so that they can calculate their own secret keys. Lain et.al [20] in 1989 was the 1st to come up with an algorithm in this approach making use of (t, n) .It consists of (k-1) members. We can also provide some papers in [2],[21],[25] with the same principle.

Coming to our paper, we are able to make a solution to this problem by providing confidentiality and authentication. We also came forward separating the insider and outsider attacks.

To achieve all the above, every user should have an account in KGS to access the group key transfer service and in turn to achieve a secret key. So, for all these transformations, we need a secret channel for message passing to all the communication entities. And also to transfer this selected group key, to all insiders of network, we need a separate and secret channel. This group key is confidential and no mathematical calculations are involved here but it is information theoretically secure.

III. Objective

Having a look at its background, we should be acquainted with: Choose two large primes p and q and calculate a public n such that $n = p*q$, which can be referred as quandary of factoring.

Practically resolving the quandary of factoring is difficult. Even though Blakely [1] and Shamir [26] developed a solution for this, it is not so efficient. According to this scheme, a whole secret key is shared among all the communication entities so that each gets a share of t. With more or equal to t shares each can calculate their secret keys. But with less than t, computation is not possible. This is called (t, n) scheme. It in turn consists of 2 algorithms:

a) Share Generation Algorithm:

Dealer D first picks a polynomial f(x) of degree (t-1) randomly: $f(x) = a_0 + a_1 x + ,.....a_{t-1} x^{t-1}$, in which the secret $S = a_0 = f(0)$ and all coefficients $a_0, a_{1,.....} a_{x-1}$ are in a finite field $IF_p = GF(p)$ with p elements.

D calculates all shares: $S_i = f(i)(\mathrm{mod}_p)$ for $i = 1,.....n$ Then,

D calculates a list of n shares $(S_i, S_2, ...S_n)$ and distributes each share S_i to corresponding shareholder P_i privately.

b) Secret Reconstruction Algorithm:

This algorithm takes any shares $(S_{i_1}, ...S_{i_t})$ as input, it can reconstruct the secret s as

$$S = f(0) = \sum_{i \in A} S_i \beta_i$$

$$= \sum_{i \in A} S_i \left(\prod_{j \in A - \{i\}} \frac{x_j}{x_{j-} x_i} \right) (\mathrm{mod}\ p),$$

$A = \{i_1, ...i_t\} \subseteq \{1, 2, ...n\}$, β_i for $i \in A$ are Lagrange coefficients. This scheme is able to satisfy all the security related issues like

a) Able to calculate the secret key only if t or more than t shares are known.
b) If not more than t shares are known, it is not able to calculate the secret key.
c) Also follows the Shamir's scheme that there are no numerical calculations and all are assume base on the above expressions.

After all this a modular inverse is to be calculated for secret reconstruction process. It is discussed in Euclid algorithm [30].

Coming to objectives the proposed protocol is distributed key generation under the consideration of

peer resource heterogeneity and security. In proposed protocol model, KGS undertakes the selection of optimized peers to participate in key generation and authenticates the peer integrity and eligibility to become part of the peer network by receiving group key. At the outset every member should register to the KGS which intern at registration selects peers with optimal resources to participate in key generation and provides those selected peers a confidential matter by which calculation of secret key is done and authenticity state of the every peer expecting to be part of the network. Then the selected peers generates group key and for each correct and authorised peer to receive group key, a checksum is appended with cipher text. All around the encryption algorithm provides this security. The confidentiality is achieved by secret sharing scheme proposed. For security, a general broadcast message is created and sent to all communication entities where its secrecy is maintained theoretically.

Considering heterogeneity of the peer resources in key generation and security is the key factor in our paper. So the primary goal is to provide security. Some important goals formulated are:

Selecting peers for key generation: Selecting peers that are optimized in terms of having resources to participate in key generation.

Fixing the key generation peer group count : The proposed protocol selects set of peers such that all other peers can receive group key from selected peer in hop level.

Key freshness : That is, the key should not be used before so that further problems may not arise.

Key Confidentiality : It is the assurance that the secret information is accessed only by authorized group members.

Key authentication : Providing authentication guarantees that generation and broadcasting of secret group key is done by KGS, a trusted organisation but not by any hackers.

In spite of all these QoS and security issues, we have 2 more threats to be worked on

a) Any hacker in person using the authenticated group user for his works done.
b) Hackers modifying the messages in their way of transfer even before reaching the destination esp. KGS.

IV. Proposed Protocol

To overcome these, the proposed protocol has 3 remedial measures.

a) Initialization of KGS.
b) User registration
c) Selecting optimal peers for Group key generation
d) Group key generation and distribution.

Initialization of KGS : In this primary step, KGS chooses optimal peers to participate in group key generation. Then KGC sends all random primes selected as shared checksums of the optimal peers to all peers participating in key generation. Then the peers selected for key generation compute n from shared checksums sent by KGC. This n is made public as stated in the proposed theory above in this paper.

User Registration: Immediately after the KGS is initialized, it is ready to use and encourages the user registrations. It in turn keeps track of all the registered users and alerts optimal peers about unauthorised peers.

Optimal peer selection for key generation : Since the heterogeneity of the peer computational resources has taken into consideration, our proposed model selects the optimal peers with eligible computational resource for group key generation.

Group key generation and distribution : As the registration phase ends with the user requests to the KGS for authentication, it sends the shared checksums of the optimal peers to all optimal peers along with the credentials of eligible peers to optimal peers selected for group key generation. Then optimal peers randomly select the secret key t of the hop level requested user and send him the message which is unique to him. By this he can access the group key.

All this transformations between the KGS and users is fallows.

Step 1: KGS receives certificates and about computational resources from Group members to initiate the key generation.

Step 2 : As the authentication, KGS responses by sending the broadcast messages to selected peers that are optimal in resources to participate in key generation.

Step3 : As a note of agreement, optimal peers send a random challenge $R_i \in Z_n^*$ to KGS.

Step 4 : KGS sends all random challenges as shared checksums of optimal nodes to all optimal nodes.

Then optimal nodes generates group key k from these shared checksums received from GKS, and generates an interpolated polynomial $f(x)$ with degree t' to pass through $(t+1)$ points, $(0,k)$ and $(x_i, y_i \oplus R_i)$, for $i = 1,2,3,....t'$. Optimal nodes also compute t additional points, OP_i for $i = 1,2,3,....t'$,on $f(x)$ and $auth = h(k, OP_1, OP_2, OP_3, ...OP_{t'})$, where h is a one-way hash function and $OP_1, OP_2, OP_3, ...OP_t$, are optimal peers. Then optimal peers send $(auth, OP_i)$, for $i = 1, ..., t'$.

Step 5: Every group member, P_i, after knowing the shared secret, $(x_i, y_i \oplus R_i)$, and other optimal

peers OP_i for $i = 1,...,|OP|$, on $f(x)$ P_i able to compute the polynomial $f(x)$ and recover the group key and then P_i computes hash value from k and OP_i for $i = 1,2,3,....t'$ then compares with $auth$ for validity.

V. Results Analysis

The experiments were conducted by developing simulation model using MXML. We build a simulation network with hops count of 80. The simulation parameters described in table 1. Authentication ensures that the buffer is properly allocated to valid packets. The simulation model aimed to compare "Authenticated Group Key Transfer Protocol Based on Secret Sharing" and proposed HTAGKTP. The performance check of these two protocols carried out against to the threats listed below.

- Rushing attack
- Denial of service
- Tunnelling
- The protection against tunnelling attack is the advantage of the HTAGKTP over AGKTP[32].

Number of nodes Range	80
Dimensions of space	1500 m × 300 m
Nominal radio range	250 m
Source–destination pairs	20
Source data pattern (each)	4 packets/second
Application data payload size	512 bytes/packet
Total application data load range	128 to 512 kbps
Raw physical link bandwidth	2 Mbps
Initial ROUTE REQUEST timeout	2 seconds
Maximum ROUTE REQUEST timeout	40 seconds
Cache size	32 routes
Cache replacement policy	FIFO
Hash length	80 bits
certificate life time	2 sec

Table1: Simulation parameters that we considered for experiments

Proposed protocols	Routing strategy	Protects from Rushing attack	Protects from Denial of service	Protects from Routing table modification	Protects from Tunneling
AGKTP[32]	P2p	Yes	Yes	No	No
HTAGKTP	P2p	Yes	Yes	Yes	Yes

Table 2 : Protocols and their ability to handle different attacks

The metrics to verify the performance of the proposed protocol are

- Data packet delivery ratio : It can be calculated as the ratio between the number of data packets that are sent by the source and the number of data packets that are received by the sink.
- Packet Delivery Fraction: It is the ratio of data packets delivered to the destinations to those generated by the sources. The PDF tells about the performance of a protocol that how successfully the packets have been delivered. Higher the value gives the better results.
- Average End To End Delay : Average end-to-end delay is an average end-to-end delay of data packets. Buffering during route discovery latency, queuing at interface queue, retransmission delays at the MAC and transfer times, may cause this delay. Once the time difference between packets sent and received was recorded, dividing the total time difference over the total number of CBR packets received gave the average end-to-end delay for the received packets. Lower the end to end delay better is the performance of the protocol.
- Packet Loss : It is defined as the difference between the number of packets sent by the source and received by the sink. In our results we have calculated packet loss at network layer as well as MAC layer. The routing protocol forwards the packet to destination if a valid route is known; otherwise it is buffered until a route is available. There are two cases when a packet is dropped: the buffer is full when the packet needs to be buffered and the time exceeds the limit when packet has been buffered. Lower is the packet loss better is the performance of the protocol.
- Routing Overhead : Routing overhead has been calculated at the MAC layer which is defined as the ratio of total number of routing packets to data packets.

Figure 1(a) shows the Packet Delivery Ratio (PDR) for basic P2P, AGKTP[32] and HTAGKTP. Based on these results it is evident that HTAGKTP recovers most of the PDR loss that observed in AGKTP[32] against to basic P2P . The approximate PDR loss recovered by HTAGKTP over AGKTP[32] is 1.5%, which is an average of all pauses. The minimum individual recovery observed is 0.18% and maximum is 2.5%. Figure 1(b) indicates AGKTP[32] advantage over

HTAGKTP in Path optimality. HTAGKTP used average 0.019 hops longer than in AGKTP[32] because of the hop level certification validation process of the HTAGKTP that eliminates nodes with invalidate certificate. Here slight advantage of AGKTP[32] over HTAGKTP can be observable.

The packet delivery fraction (PDF) can be expressed as:

$$P' = \sum_{f=1}^{e} \frac{R_f}{N_f}$$

$$P = \frac{1}{c} * P'$$

- P is the fraction of successfully delivered packets,
- c is the total number of flow or connections,
- f is the unique flow id serving as index,
- R_f is the count of packets received from flow f
- N_f is the count of packets transmitted to flow f .

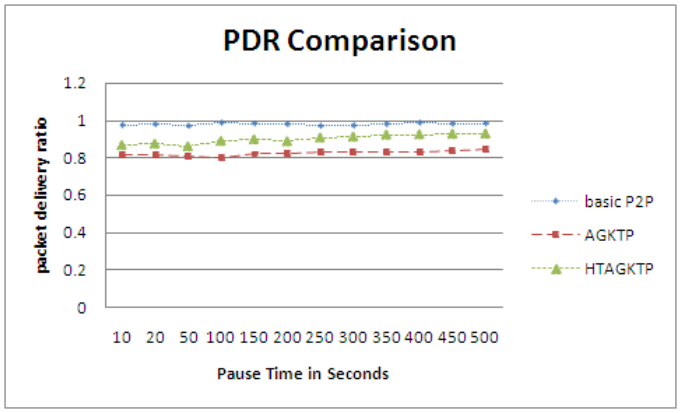
Figure 1(c) confirms that HTAGKTP is having fewer packets overhead over AGKTP[32]. Due to stable paths with no compromised or victimized nodes determined by HTAGKTP this advantage become possible. The Packet overhead observed in AGKTP[32] is average 5.29% more than packet overhead observed in HTAGKTP. The minimum and maximum packet overhead in AGKTP[32] over HTAGKTP observed is 3.61% and 7.29% respectively.

MAC load overhead is slightly more in HTAGKTP over AGKTP[32]. We can observe this in Figure 1(d), which is because of additional control packet exchange in HTAGKTP for neighbour hop validation through certificate exchange. The average MAC load overhead in HTAGKTP over AGKTP[32] 1.64%. The minimum and maximum MAC load overhead observed is 0.81 and 3.24% respectively.

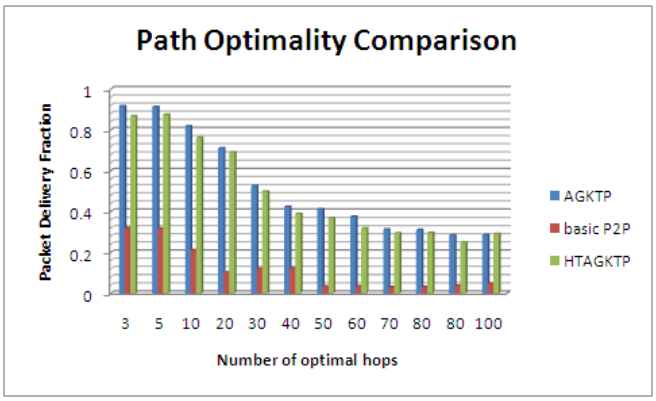
In all these evaluation strategies the results derived for basic P2P are interesting. In all metrics except path optimality, basic P2P performed well since it is not considering any security issue as routing parameter, and it is delivering better QOS under no security threat in routing assumption, which is not true in real time practices. In path optimality validation among three considered protocols basic P2P stands last because it is not considering any security constraints, hence identifies unstable paths.

VI. Conclusion

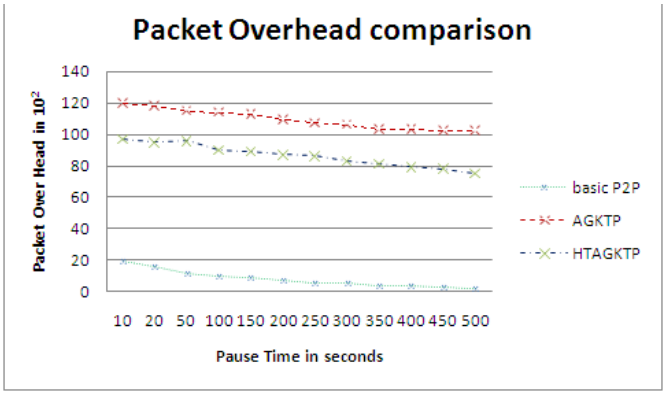
Tight security mechanisms are needed to allow secure communication among the group members. Thus, a communication session must have security services to provide authentication, integrity, and confidentiality. Group Key (GK) is the primary and key part of the safe group communication. The performance of GK generation process, which is required for secure communication, may degrade due to less performing members. Thus, the generation process must be done is a more precise way but filtering less performing members. Many changes are occurring in the recent years as increase in usage of mobile computers, network clusters communication with standard servers. Apart from this, heterogeneity and distributed computer environment became common in the current internet world. Thus, GK management system must consider various parameters, differences and environments involved in the communication. These considerations as the basis, the effectiveness of HTAGKTP protocol in comparison to AGKTP[32] is proved. This protocol improves the efficiency by considering the parameters effecting the performance i.e. computational delay and network latency. Thus, this research is aimed at and thus proved that GKGP is more efficient and maximizes the applicability of communication.



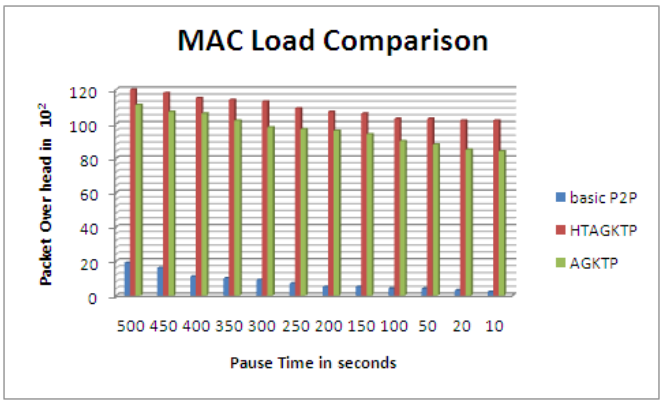
(a) Packet delivery ratio comparison using line chart



(b) Bar chart representation of Path optimality



(c) A line chart representation of Packet overhead comparison report



(d) Mac load comparison represented in bar chart format

Figure 1: Evaluation report of HTAGKTP performance over AGKTP[32]

REFERENCES RÉFÉRENCES REFERENCIAS

1. G.R. Blakley, "Safeguarding Cryptographic Keys," Proc. Am. Federation of Information Processing Soc. (AFIPS '79) Nat'l Computer Conf., vol. 48, pp. 313- 317, 1979.
2. S. Berkovits, "How to Broadcast a Secret," Proc. Eurocrypt '91 Workshop Advances in Cryptology, pp. 536-541, 1991.
3. R. Blom, "An Optimal Class of Symmetric Key Generation Systems," Proc. Eurocrypt '84 Workshop Advances in Cryptology, pp. 335-338, 1984.
4. C. Blundo, A. De Santis, A. Herzberg, S. Kutten, U. Vaccaro, and M. Yung, "Perfectly Secure Key Distribution for Dynamic Conferences," Information and Computation, vol. 146, no. 1, pp. 1-23, Oct. 1998.
5. C. Boyd, "On Key Agreement and Conference Key Agreement," Proc. Second Australasian Conf. Information Security and Privacy (ACISP '97), pp. 294-302, 1997.
6. E. Bresson, O. Chevassut, D. Pointcheval, and J.-J. Quisquater, "Provably Authenticated Group Diffie-Hellman Key Exchange," Proc. ACM Conf. Computer and Comm. Security (CCS '01), pp. 255-264, 2001.
7. J.M. Bohli, "A Framework for Robust Group Key Agreement," Proc. Int'l Conf. Computational Science and Applications (ICCSA '06), pp. 355-364, 2006.
8. M. Burmester and Y.G. Desmedt, "A Secure and Efficient Conference Key Distribution System," Proc. Eurocrypt '94 Workshop Advances in Cryptology, pp. 275-286, 1994.
9. R. Canetti, J. Garay, G. Itkis, D. Micciancio, M. Naor, and B. Pinkas, "Multicast Security: A Taxonomy and Some Efficient Constructions," Proc. IEEE INFOCOM '99, vol. 2, pp. 708-716, 1999.
10. J.C. Cheng and C.S. Laih, "Conference Key Agreement Protocol with Non Interactive Fault-Tolerance Over Broadcast Network," Int'l J. Information Security, vol. 8, no. 1, pp. 37-48, 2009.
11. W. Diffie and M.E. Hellman, "New Directions in Cryptography," IEEE Trans. Information Theory, vol. IT-22, no. 6, pp. 644-654, Nov. 1976.
12. M. Eltoweissy, M.H. Heydari, L. Morales, and I.H. Sudborough, "Combinatorial Optimization of Group Key Management," J. Network and Systems Management, vol. 12, no. 1, pp. 33-50, 2004.
13. A. Fiat and M. Naor, "Broadcast Encryption," Proc. 13th Ann. Int'l Cryptology Conf. Advances in Cryptology (Crypto '93), pp. 480-491, 1994.

14. H. Harney, C. Muckenhirn, and T. Rivers, "Group Key Management Protocol (GKMP) Architecture," RFC 2094, July 1997.
15. K.H. Huang, Y.F. Chung, H.H. Lee, F. Lai, and T.S. Chen, "A Conference Key Agreement Protocol with Fault-Tolerant Capability," Computer Standards and Interfaces, vol. 31, pp. 401-405, Jan. 2009.
16. IEEE 802.11i-2004: Amendment 6: Medium Access Control (MAC) Security Enhancements, 2004.
17. I. Ingemarsson, D.T. Tang, and C.K. Wong, "A Conference Key Distribution System," IEEE Trans. Information Theory, vol. IT-28, no. 5, pp. 714-720, Sept. 1982.
18. J. Katz and M. Yung, "Scalable Protocols for Authenticated Group Key Exchange," J. Cryptology, vol. 20, pp. 85-113, 2007.
19. C. Lain, J. Lee, and L. Harn, "A New Threshold Scheme and Its Application in Designing the Conference Key Distribution Cryptosystem," Information Processing Letters, vol. 32, pp. 95-99, 1989.
20. C.H. Li and J. Pieprzyk, "Conference Key Agreement from Secret Sharing," Proc. Fourth Australasian Conf. Information Security and Privacy (ACISP '99), pp. 64-76, 1999.
21. A. Perrig, D. Song, and J.D. Tygar, "Elk, A New Protocol for Efficient Large- Group Key Distribution," Proc. IEEE Symp. Security and Privacy, pp. 247-262, 2001.
22. G. Saze, "Generation of Key Predistribution Schemes Using Secret Sharing Schemes," Discrete Applied Math., vol. 128, pp. 239-249, 2003.
23. A. Shamir, "How to Share a Secret," Comm. ACM, vol. 22, no. 11, pp. 612- 613, 1979.
24. A.T. Sherman and D.A. McGrew, "Key Establishment in Large Dynamic Groups Using One-Way Function Trees," IEEE Trans. Software Eng., vol. 29, no. 5, pp. 444-458, May 2003.
25. D.G. Steer, L. Strawczynski, W. Diffie, and M.J. Wiener, "A Secure Audio Teleconference System," Proc. Eighth Ann. Int'l Cryptology Conf. Advances in Cryptology (Crypto '88), pp. 520-528, 1988.
26. M. Steiner, G. Tsudik, and M. Waidner, "Diffie-Hellman Key Distribution Extended to Group Communication," Proc. Third ACM Conf. Computer and Comm. Security (CCS '96), pp. 31-37, 1996.
27. D.R. Stinson, Cryptography Theory and Practice, second ed., CRC Press, 2002.
28. W.G. Tzeng, "A Secure Fault-Tolerant Conference Key Agreement Protocol," IEEE Trans. Computers, vol. 51, no. 4, pp. 373-379, Apr. 2002.
29. Harn, L.; Changlu Lin; , "Authenticated Group Key Transfer Protocol Based on Secret Sharing," Computers, IEEE Transactions on , vol.59, no.6, pp.842-846, June 2010, doi: 10.1109/TC.2010.40

This page is intentionally left blank

Global Journal of Computer Science and Technology
Volume 12 Issue 7 Version 1.0 April 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Reverse Auction in Pricing Model

By Nagabhushan.S.V, Praveen kulkarni, Dr.K.N. Subramanya & Dr.G.N.Srinivasan
RV College of Engineering

Abstract - Dynamic price discrimination adjusts prices based on the option value of future sales, which varies with time and units available. This paper surveys the theoretical literature on dynamic price discrimination, and confronts the theories with new data from airline pricing behavior, Consider a multiple booking class airline-seat inventory control problem that relates to either a single flight leg or to multiple flight legs. During the time before the flight, the airline may face the problems of (1) what are the suitable prices for the opened booking classes, and (2) when to close those opened booking classes. This work deals with these two problems by only using the pricing policy. In this paper, a dynamic pricing model is developed in which the demand for tickets is modeled as a discrete time stochastic process. An important result of this work is that the strategy for the ticket booking policy can be reduced to sets of critical decision periods, which eliminates the need for large amounts of data storage.

Keywords : *Bidding, dynamic pricing, electronic markets, group-buying discounts, Internet based selling, market microstructure, online retailing, pricing mechanism.*

GJCST Classification: *K.4.4*

Strictly as per the compliance and regulations of:

Reverse Auction in Pricing Model

Nagabhushan.S.V [α], Praveen kulkarni [α], Dr.K.N. Subramanya [σ] & Dr.G.N.Srinivasan [σ]

Abstract - Dynamic price discrimination adjusts prices based on the option value of future sales, which varies with time and units available. This paper surveys the theoretical literature on dynamic price discrimination, and confronts the theories with new data from airline pricing behavior, Consider a multiple booking class airline-seat inventory control problem that relates to either a single flight leg or to multiple flight legs. During the time before the flight, the airline may face the problems of (1) what are the suitable prices for the opened booking classes, and (2) when to close those opened booking classes. This work deals with these two problems by only using the pricing policy. In this paper, a dynamic pricing model is developed in which the demand for tickets is modeled as a discrete time stochastic process. An important result of this work is that the strategy for the ticket booking policy can be reduced to sets of critical decision periods, which eliminates the need for large amounts of data storage.

Keywords : Bidding, dynamic pricing, electronic markets, group-buying discounts, Internet based selling, market microstructure, online retailing, pricing mechanism.

I. Introduction

In Dynamic pricing mechanisms where buyers and sellers actively engage in the price discovery process. Today, traditional dynamic pricing mechanisms serve as the core business models for many Internet-based electronic markets. Indicative of its attractiveness to consumers in the Internet marketplace, the registered users of eBay increased from 1.2 million to 2.1 million during the last three months and about $750 Million worth transactions were conducted on their website that year. Other new approaches like online trading and online auctions are also attracting consumers, however. They include dynamic Pricing mechanisms such as the name-your-own-price.E-bay is the world's largest auction site where the buyers and seller are to be participating in the auction mechanism-bay is one of the most profitable e-business site, in one day 5,00,000 items may added in the auction site, this strategy can work throughout at the end of the day. The initial model of c2c auctions entirely managed by the technology, buyers can easily buy the products at any time in the auction site. Some sites use dynamic pricing in auction mechanism and they are like priceline.com this site involve the dynamic variations of the prices in the market. The other site which is known as travelbids.com in which the buyers can bid their money for travelling expenses through auction mechanism. There are some websites like Priceline.com, TravelBids.com and BuyersEdge.com which uses dynamic pricing mechanism. Retailer itself, As this group-buying discount pricing mechanism is still a new phenomenon on the Internet, there have been no studies that have examined the performance of this market microstructure, and the nature of the bidder behaviour that can be observed as the market operates. Research on consumer behaviour under the group-buying market structure can help both academic researchers and industrial practitioners better understand this new kind of market intermediary, and the efficacy of the market mechanism that it provides to market participants.

a) Basics of Auction Mechanism

The commonly used basic auction mechanisms can be classified into the following two main categories based on the number of bidding sides: a single auction and a double Auction , In a single auction, participants can take part only in one side of an auction (i.e., be either an auctioneer or a bidder). In a double auction, participants are free to take part in both side of an auction. Other kinds of auction mechanism are as follows.

Single auction mechanism: This type of auction can be divided into an open-outcry auction and a sealed bid auction based on bidding methods. In an open-outcry auction, the bids are open to public and Bidders can adjust their bids in the full knowledge of other bids. In a sealed bid auction, only a bidder and the auctioneer can communicate with each other, and bidder to bidder Communication is forbidden. The basic single auction sub-types with open-cry bid are English and Dutch auctions: English auction: In an English auction, an auctioneer creates an auction market and proceeds to solicit in open successively higher bids from the bidders until no one raises the bid. The highest bidder is the winner and pays the price he/she bid. Dutch auction: An auctioneer announces the bids to all bidders. The auctioneer starts the bidding at an extremely high price and then progressively lowers it until a buyer claims an item by calling "mine", or by pressing a button that stops an automatic clock. The winner pays the price bid at the stop time.

Double auction mechanism: A double auction admits multiple buyers and multiple sellers concurrently into the market. Thus, the double auction must match bids of the both sides in the market. The double auction

Author α : Department of MCA,BMSIT, 6thsem MCA, Asst.Prof Nagabhushan.S.V1. E-mail : kk_praveenkulkarni@rediffmail.com, bhushan.svn@gmail.com
Author σ : Professor : Dept of IEM, RV College of Engineering. E-mail : sansa96@gmail.com
Author σ : Department of Information Science,RVCE. E-mail : srinivasangn@rvce.edu.in

can be divided into two main classes based on the auction clearing time: Call Market and Continuous Double Auction (CDA). In a Call Market, bids are collected over a specific time interval from both sellers and buyers in a sealed manner.

b) Extended Types of Auction Mechanism

Multi-attribute Auction (MA): A Multi-attribute Auction allows bidders to bid on various attributes beyond the price. In this type of an auction, the auctioneer selects winners based on the price as well as on those various attributes. Thus, the overall utility of a deal for the buyer must consider not only the price of the auctioning item, but also a combination of the different attributes, Combinatorial Auction (CA):In a Combinatorial Auction, each bidder offers a bid for a collection of resources (of the bidder's choosing) rather than placing a bid on each item separately. This enables the bidder to express dependencies and complementarities between various resources. The auctioneer selects such set of these combinatorial bids that result in the highest revenue without assigning any item to more than one bidder. Snadholm showed that the number of possible allocations in a combinatorial auction is where m is the number of items traded in the auction.

Generalized Vickrey Auction (GVA): Another mechanism for determining prices for an allocation of multiple units of resources is the Generalized Vickrey Auction. In GVA, the price of a bidder k in the efficient allocation is computed by deducting the sum of payments of all other bidders in an allocation from the sum of the payments that would be obtained from those bidders in the optimum allocation where the bidder k removed from the allocation.

Finally, enough bidders must be willing to participate in future auction roundsto prevent a collapse of market prices. Because of these requirements, previously designedauctions cannot work efficiently in e-service markets.A static time differential pricing mechanism in which two or more tiers ofon/off peak rates are used can improve efficiency by partially matching lower (higher) demand with lower (higher) price. However, this mechanismstill remains inflexible, since demands of buyers do not follow a step function, but rather gradually shift from on- to off-peaks and back. Dynamic pricing mechanism that adapts tochanging market conditions constantly is more efficient. It maintains highresource utilization and the seller's revenue in variety of market conditions.

II. Literature Review

In the year 2009, the paper proposed by Roumen Vragov, Di Shang, Karl R. Lang, In the paper "Auction based E-Procurement Mechanism Design ", They sugest that a view of an emerging phenomenon The E-procurement [3]. But the drawback of a new mechanism is designed based on E-auction mechanism; one problem occurred during this paper that without internet we are not able to implement the auction strategies.

In the year 2009,the paper proposed by Hila Etzion and Scott Moore ,In the paper "Secure keyword auction:[4] preserving privacy of bidding prices and CTRs",they suggest that we use a simulation model to extend previous analytical research on a firm selling consumer goods online using posted price and auction at the same time.Three research streams are useful in informing our how their expectations about future prices develop, and how their attitudes toward risk affect their behavior.but the drawback in this paper[16] is Bidder and buyer reactions to prices are especially worthwhile to consider in the context of group buying discount electronic markets. Even though Drakopoulos doesn't use the term explicitly,[14] he explains the main idea of price indifference thresholds, which refer to the minimum price changes required for consumers to detect the differences. By incorporating a threshold into a demand curve, he proposes that when a price change is smaller than some critical threshold, consumers will not detect the price change, resulting in no change in demand. In our research, we will refer to the related term, price threshold, to indicate proximity in order quantity terms to the quantity-price combination that reflects a drop in price to the lower-tier[2].

In the year 2009,the paper proposed by Robert J. Kauffman, in the paper "New Buyers' Arrival under Dynamic Pricing Market Microstructure" they suggest that Using data collected from MobShop.com, the current study tested a model that predicts bidder participation and order arrivals in an electronic market that uses an innovative group-buying discount market microstructure.[4][5]The problem can be occurring as some buyer's donot use the electronic equipments in the market. The variations of a cost can be high, the consumer can pay high amount for new features and new items can be added.

In the year 2009, the paper proposed by Shamik Sengupta and Mainak Chatterjee, in the paper "An Economic Framework for Dynamic Spectrum Access and Service Pricing" they suggest that Dynamic spectrum allocation coupled with fine granularity switching of services by end-users will engender a flexible and competitive environment for trading wireless services.[6][8] The drawback of this paper to implement the radio signals and their dynamic mechanism.In the year 2010,the paper proposed by Wei-Yu Lin, Guan-Yu Lin, Hung-Yu Wei, In the paper "Dynamic Auction Mechanism for resource allocation" they suggest The main contribution of this paper is developing a new resource allocation algorithm by applying auction method into the resource allocation, but one problem occurred while in the allocation of resources in the market, it is not able to determine the approximate profit

and loss for a product in market.In the year 2011, the paper[9] proposed by lin gao, xinbing wang and youyunxu, in the paper "Multiauctioneer Progressive Auction for Dynamic Spectrum Access", they suggest the problem of residual channel allocation We propose an auction-based mechanism MAP, in which each PO systematically raises the trading prices and each SU subsequently decides whether to buy a spectrum band and from which PO he is going to buy a spectrum band.

In the year 2011, the paper proposed by Andrew Chang, in the paper[8] "Time dynamics of user behavior in a series of overlapping electronic mechanism" they suggest shed new lights on time dynamics of online user behavior in the market which comprise a series of multiple overlapping auctions, the error can be occur during the series of electronic equipments there might be a loss of data and loss of important database.

In the year 2011, the paper proposed by R. J. Thomas, T. D. Mount, R. Zimmerman, W. D. Schulze, in the paper[11] "Testing the Effects of Price Responsive Demand on Uniform Price and Soft-Cap Electricity Auctions", they suggest that the effectiveness of a price while bidding in a market, the price can be low or high during auction mechanism. The drawback of this paper is the user can know about variations of prices in a market. But one problem occurs during researching of this paper that the user not able to extra features or more items.

In the year 2011, the paper proposed by Shabnam Sodagari, Alireza Attar, in the paper "On a Truthful Mechanism for Expiring Spectrum Sharing in Cognitive Radio Networks" they suggest a powerful tool to realize such subletting of Spectrum is auction mechanisms. In this paper, we proposed a dynamic, online auction for secondary spectrum access, the drawback of this paper is proposed auction structure, secondary CRs not only will submit their valuation of auctioned spectrum band, but also their arrival and departure time instances.

Computerized reservation systems were developed in the 1950s to keep track of airline seat booking and fare information. Initially these were internal systems, but were soon made available to travel agents. Deregulation of airline pricing in 1978 permitted much more extensive use of the systems for economic activity[12], especially pricing. The initial development of dynamically adjusted pricing is often credited to American Airlines' Robert Crandall, as a response to the rise of discount airline People's express in the early 1980s.

a) Problem Formulation

The complexity and opaqueness of airline pricing has grown over time. As a result, the "yield management"[12]systems employed by airlines for pricing have become one of the most arcane and complex information systems on the planet, and one with a very large economic component. Airline pricing represents a great challenge for modern economic analysis because it is so distant from the "law of one price" level of analysis. This surveys the theoretical literature, which is mostly found in operations research journals, develops some new theory, assesses the holes in our knowledge, and describes some results from a new database of airline prices. Dynamic pricing, which is also known as yield management or revenue management, is a set of pricing strategies aimed at increasing profits. The techniques are most useful when two product characteristics co-exist. First, the product expires at a point in time, like hotel rooms, airline flights, generated electricity, or time-dated ("sell before") products. Second, capacity is fixed well in advance and can be augmented only at a relatively high marginal cost. These characteristics create the potential for very large swings in the opportunity cost of sale, because the opportunity cost of sale is a potential foregone subsequent sale. The value of a unit in a shortage situation is the highest value of an unserved customer. Forecasting this value given current sales and available capacity represents dynamic pricing.

III. Empirical Solution

The dynamic pricing in auction mechanism involve some existing sytems like e-bay auction technology site and multi-agent wireless system for dynamic and local combined pricing, allocation and billing. The e-bay is an world's largest auction site where this site shows the interaction between the seller and buyer-bay is one of the profitable e-business site, on e-bay people can buy and sell just about anything, the company collects a submission see upfront and a commission as a percentage of the sale amount.

Example1: The auction process in e-bay: The seller fills in the appropriate registration information post the description of the item for sale specifying a minimum opening bid, if a successful bid is made ,the seller and the buyer negotiate the payment method, shipping details, warranty and other perticulers,e-bay is the interface through which sellers and buyers can conduct business. In 2011 e-bay started to auction fine art in collaboration with icollector.com of the united kingdom and sothebys.com, e-bay operates globally buyers from more than 150 other countries participate, trading can be made from anywhere and at any time.

Example 2 : Multi-Agent Wireless System for Dynamic and Local Combined Pricing, Allocation and Billing: In established communications systems with an underlying provider infrastructure the market is designed according to the Fixed Price Model (FPM). A user can get access to the network only if there are free resources controlled by the Base Station (BS) within a cell. Furthermore, he has to accept the fixed price for

awide area and quasi-static in time. The user has to pay the same price whether there is a high or a low demanding the cell. Thus, if his preferences and purchase power allow him to spend more money for using Radio Resource Goods (RRG), he is not able to influence the allocation. For the same reason the operator misses the chance to increase his monetary gain by adapting the price for RRG to the users RRG evaluation. To overcome these problems will introduce a distributed, dynamical and combined pricing, allocation and billing system, suitable for wireless infrastructure communications systems which are capable to manage multi-homing. By applying Cognitive Radio (CR) abilities not only to the allocation but also to this combined architecture, it is mandatory to dynamically allocate RRG by an Auction Sequence (AS) to exploit the CR abilities the repetition of auctions should happen very fast up to milliseconds to react on the load dynamic. A class of auctions, the multi-unit sealed-bid auction, is suitable to execute the auction within specified time and with an acceptable signalling effort in comparison to sequential auctions.The economical field auction theory mainly deals with single shot auction, but not with auction repetition. The framework of the optimal multi-unit auction is proposed in, but does not deal with repetitions and does not include learning facilities of the UTs. Gaining experience of the past and applying it in the current auction destroy the independence of the single shot auctions.the asymmetric balance of negotiation power needs to be considered in auction design. The prices bid in a basic auction are dependent only on the bidder's willingness to pay. This means that intentions of only bidders, but not the auctioneer, are reflected in the auction winning prices. To restore the symmetric balance of negotiating power, the reservation price (RPA) and cancelable (CA) auctions were proposed. In RPA, only bids higher than the auctioneer's reservation price are considered during winner selection. On the other hand, in CA, if the resulting revenue of an auction round does not meet the minimum requirement of the auctioneer, the entire auction round is cancelled. By providing reservation price or cancellation option to the auctioneer, the asymmetric negotiation power problem is resolved. However, when the perishable resources are traded, both of these auctions cause resource waste. In RPA, the reservation price restricts the number of winners. Hence, the resources unused because of this restriction are wasted. In CA, the cancellation of an auction round wastes the entire stock of resources that are allocated to this auction round.

IV. Conclusion

1. It is very clear that advances in internet and e-commerce technologies have opened up rich opportunities for reaping the benefits of dynamic pricing. Companies resorting to dynamic pricing strategies are increasing in number steadily. Moreover, increasingly complex dynamic pricing strategies are being tried out. In this paper, we have covered the following topics.

We have shown that the fixed pricing paradigm is giving way to a dynamic pricing paradigm in e-business markets and those dynamic pricing strategies: Conditions under which dynamic pricing strategies will outperform fixed pricing strategies have been enunciated.

Searching for more efficient ways of deciding the auctioneer's optimal bid that defines the DW and PW classes. Seeking a better way to find the optimal winning score coefficient for the winner selection strategy. Attempting to identify more efficient and simpler pricing rules that preserve incentive compatibility but achieve higher auctioneer's revenue

Software modelling of dynamic pricing mechanism which illustrates the reliability and efficiency of the proposed model.

Acknowledgement

We thank all the authors for the information support

V. Future Works

We can develop a pricing model for substitutable flights where customers choose among the available flights. To overcome computational problems posed by the formulation's multi-dimensional state and action spaces, we considered heuristics based on pooling ideas. We also derived easily-computable separable bounds for the value function of our model. Policies motivated by these bounds were shown numerically to be near optimal for a range of problem instances, and to dominate the policies from pooling in most cases. Our results suggest that pooling heuristics perform well for symmetric problems in which (a) customers, when viewed as a population, are mostly indifferent in their preferences over flights and (b) the flights have the same seating capacity. However, the pooling heuristics can perform poorly for asymmetric problems. The approaches motivated by the separable bounds do not suffer from such shortcomings, and remain implementable for large problems.

References

1. Lin GAO, Xinbing Wang, Youyun Xu and the prepared paper title was "Multi-auctioneer Progressive Auction for Dynamic Spectrum Access" in 2011.
2. Wei-Yu Lin ,Guan-Yu Lin, Hung-Yu Wei and the prepared paper title was" Dynamic Auction Mechanism for Resource Allocation" in 2010
3. Roumen Vragov, Di Shang, Karl R. Lang and the prepared paper title was" Auction based E-Procurement Mechanism Design" in 2009

4. Hila Etzion and Scott Moore and the prepared paper title was" Secure keyword auction: preserving privacy of bidding prices and CTRs" in 2009
5. Robert J. Kauffman and bin wang and the prepared paper title was "New Buyers' Arrival under Dynamic Pricing Market Microstructure"in2010
6. Roumen Vragov, Di Shang, Karl R. Lang and the prepared paper title was "Should Online Auctions Employ Dynamic Buyout Pricing Models "in 2009
7. Hila Etzion and Scott Moore and the prepared paper title was "Simulation of Online Selling with Posted-price and Auctions: Comparison of Dual Channel's Performance under Different Auction Mechanisms"in 2008.
8. Seokjoo Andrew Chang and the prepared paper title was "time dynamic pricing for user behavior" in 2010.
9. Andrews, W. "The New Laws of Dynamic Pricing".Internet World, Dec 15, 2009
10. Bakos, Y. "The Emerging Role of Electronic Marketplaces on the Internet". Communications of the ACM, Aug 2008.
11. Chang, R. P., Hsu, S. T., Huang, N. and Rhee, S. G. "The Effects of Trading Methods on Volatility and Liquidity: Evidence From the Taiwan Stock Exchange". Journal of Business Finance and Accounting, Jan/Mar 2010
12. Keegan, P. "Online Auctions: From Seedy Flea Markets to Big Business". *Upside,* Jul 2009, 11(7), U.S. edition, 70-81.
13. X.-Y. Li, P. Xu, S. Tang, and X. Chu, "Spectrum Bidding in Wireless Networks and Related," Proc. Conf. Computing and Combinatorics (COCOON '08), 2008.
14. Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, "Enabling public verifiability and data dynamics for storage security in cloud computing," *in Proc. of ESORICS'09, Saint Malo, France, Sep. 2009*.
15. X. Zhou, S. Gandhi, S. Suri, and H. Zheng, "eBay in the Sky: Strategy-Proof Wireless Spectrum Auctions," Proc. ACM MobiCom, Sept. 2008.
16. X. Zhou and H. Zheng, "TRUST: A General Framework for Truthful Double Spectrum Auctions," Proc. IEEE INFOCOM, 2009.
17. J. Lian, Y. Liu, K. Naik, and L. Chen, "Virtual Surrounding Face Recasting with Guaranteed Message Delivery for Ad Hoc and Sensor Networks," IEEE/ACM Trans. Networking, vol. 17, no. 1,pp. 200-211, Feb. 2009.
18. G. Tesauro and J. O. Kephart. Pricing in agent economies using neural networks and multiagent q-learning. In *Proc. Workshop ABS-3:* Learning about, from and with other agents (held in conjunction with IJCAI '99), Stockholm, 2009
19. B. K. Szymanski, and J. S. Lee, "Return On Investment Constraints on Bidding and Revenue of Sponsored Search Auction", Submitted to INFORMS Journal on Computing, July 2010.
20. Simon Board*y* Andrzej Skrzypacz on "Revenue Management with Forward-Looking Buyers "submitted on nov 2010.

Global Journal of Computer Science and Technology Volume XII Issue VII Version I April 2012
38

This page is intentionally left blank

Global Journal of Computer Science and Technology
Volume 12 Issue 7 Version 1.0 April 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Comparison Study and Review on Object- Oriented Metrics

By Sanjay Kumar Dubey, Amit Sharma & Dr. Ajay Rana
Amity university

Abstract - The best elucidations to software development problems are regularly touted as object-oriented processes. The popularity of object-oriented design metrics is essential in software engineering for measuring the software complexity, estimating size, quality and project efforts. There are various approaches through which we can find the software cost estimation and predicates on various kinds of deliverable items. Object-oriented metrics assures to reduce cost and the maintenance effort by serving as early predictors to estimate software faults. Such an early quantification augments the quality of the final software. This paper reviews object-oriented metrics. A comparison table is maintained via which we can analyze the difference between all the object-oriented metrics effectively.

Keywords : *Object-Oriented, class, attributes.*

GJCST Classification: *D.2.0*

Strictly as per the compliance and regulations of:

Comparison Study and Review on Object-Oriented Metrics

Sanjay Kumar Dubey [α], Amit Sharma [σ] & Dr. Ajay Rana [ρ]

Abstract - The best elucidations to software development problems are regularly touted as object-oriented processes. The popularity of object-oriented design metrics is essential in software engineering for measuring the software complexity, estimating size, quality and project efforts. There are various approaches through which we can find the software cost estimation and predicates on various kinds of deliverable items. Object-oriented metrics assures to reduce cost and the maintenance effort by serving as early predictors to estimate software faults. Such an early quantification augments the quality of the final software. This paper reviews object-oriented metrics. A comparison table is maintained via which we can analyze the difference between all the object-oriented metrics effectively.

Keywords : *Object-Oriented, class, attributes.*

I. Introduction

Object-Oriented design is more beneficial in software development environment and object-oriented design metrics is an essential feature to measure software quality over the environment [1]. object-oriented design is those design which contained all the properties and quality of software that is related to any large or small project [2]. It is a degree through which a system object can hold a particular attribute or characteristics. object-oriented is a classifying approach that is capable to classify the problem in terms of object and it may provide many paybacks on reliability, adaptability, reusability and decomposition of problem into easily under stood objects and providing some future modifications [3]. Software metrics makes it possible for software engineer to measure and predict software necessary resource for a project and project work product relevant to the software development effort. Metrics provide insight necessary to create and design model through the test. It also provide a quantative way to access the quality of internal attributes of the product, thereby it enables the software engineer to access quality before the product is build [4]. Metrics are the crucial source of information through which a software developer takes a decision for design good software. Some metrics may be transformed to serve their purpose for a new environment.

Author α σ ρ : Department of Computer Science & Engineering Amity School of Engineering and Technology Amity University, Sec-125, NOIDA, (U.P.), India.
E-mail α : skdubey1@amity.edu
E-mail σ :amit.krsharma123@gmail.com
E-mail ρ :ajay_rana@amity.edu

II. Literature Survey

Abreu *et al.* [37] provides a new classification framework for the TAPROOT. This framework was defined with the other two independent vectors these are category and granularity. Six categories of Object-Oriented metrics were defined are design metrics, complexity metrics, size metrics, quality metrics, productivity metrics and reuse metrics and also proposed three Levels of granularity are software, class and methods but no empirical/theoretical base for the metrics was provided.

M. Alshayeb *et al.* [13] have given two iterative procedures for the pragmatic study of object oriented metrics. They include the short-cycled agile process and the long cycled framework evolution process. By observing the results, it can be seen that the design efforts and source lines of code added, changed, and deleted were triumphantly predicted by object-oriented metrics in short-cycled agile process where as in the case long-cycled framework process the same features were not successfully predicted by it. This has shown that the design and implementation changes during development iterations can be predicted by object-oriented Metrics, but the same cannot be the case with long-term development of an established system. R.D.Neal *et al.* [20] also gives the study for the validation of the object-oriented software metrics and found that some of the proposed metrics could not be considered as the valid measure for the dimension then, they could be measured. He defined a model based on measurement theory of the validation through which they can proposed 10 new metrics – Potential Methods Inherited (PMI), Proportion of Methods Inherited by a Subclass (PMIS), Density of Methodological Cohesiveness (DMC), Messages and Arguments (MAA), Density of Abstract Classes (DAC), Proportion of Overriding Methods in a Subclass (POM), Unnecessary Coupling through Global Usage (UCGU), Degree of Coupling Between Class Objects (DCBO), Number of Private Instance Methods (PrIM), and Strings of Message Links (SML).

R. Harrison *et al.* [12]suggested a statistical model which is obtained from the logistic regression for identifying threshold values for the Chidamber and Kemerer metrics. The process is authenticated empirically on a large Open-Source System- the Eclipse project. Their conclusion depending on the experimental

results is that the Chidamber and Kemerer Metrics have threshold effects at different risk levels. The usefulness of these thresholds on later releases was authenticated with the aid of decision trees. Another conclusion by L.H. Ethzkorn [23] is that the chosen threshold values were more precise than those were chosen depending on either intuitive perspectives or on data distribution parameters. object-oriented design metrics has also been assign the high level design quality attributes for the object-oriented software with the help of hierarchical model. H. Lieu. *et al.* [33] have given perception that quality of software also plays an important role in terms of safety aspects and financial aspects. They bridged the gap between quality measurement and design of these metrics, with the help of measuring the excellence of object-oriented designs during development and re-development process of the software.

M. Subramanyam *et al.* [34] proposed some Metrics suites and concluded that for the developers, designs metrics are very important to know the design aspects of the software and to enhance the quality of software. Rachel Harrison *et al.* [35] discussed about the six properties of metrics for object-oriented design (MOOD) Metrics and measured the object-oriented features like Inheritance, coupling, encapsulation, and polymorphism. In the result they showed that the metrics could be used to provide an overall assessment of the system. A. Goldberg *et al.* [46] have experimentally checked size estimation models that are object-oriented. The pragmatic examination of object-oriented Function Points has been extended to a considerable amount with the aid of a bigger data set and by comparing Object Oriented Function Points with other predictors of LOC (Lines of Code) in their work. Linear models where the independent variable is either a conventional Object-Oriented entity or an Object-Oriented Function Points-related measure were built and assayed by using a cross validation approach. C. Shyam *et al.* [14] suggests some software metrics through which we can calculating the quality of modularization of an object oriented software. They aimed that it provide a set of metrics for the large scale object oriented software system with having some dependencies and also provide some metrics for characterizing the quality for modularization regarding the APIs of the one side. On the another side, they provide some object-oriented dependencies like inheritance, associates relationship and base class designing. Y. Zhou *et al.* [54] considered the fault severity using the machine learning methods with their experimental exploration of fault proneness which predict the capability of object-oriented design metrics and all of these of the predictions and the fault severity are also taken from the domain NASA data sets. J. Xu. *et al.* [53] have proposed an object-oriented metrics which describes the fault estimation using empirical analysis and also used the CK metrics to apprise the number of faults in the particular program. This also includes some neural and fuzzy technique. At last, the result showed that we can get a dependable fault by using CBO, RFC, WMC, SLOC. Here SLOc is more considerable for the effect on the number of defects. C. Neelamegan *et al.* [45] surveyed four object-oriented metrics and mostly focused on the measurements that are totally applied on the design and class characteristics. Dr. B.R. Sastry *et al.* [42] trying to implement the graphics user interaction with the aid of software metrics and also tried to determine the quantity and quality of object oriented software development lifecycle.

III. Review of Metrics

Some object-oriented metrics for the object-oriented software development. These metrics are-

A. Chen Metrics
B. Morris's Metrics
C. Lorenz and Kidd Metrics
D. MOOSE Metrics
E. EMOOSE
F. MOOD Metrics
G. Goal Question Metrics
H. QMOOD Metrics
I. LI Metrics
J. SATC for object oriented metrics

a) Chen Metrics

Chen *et al.* [30] proposed software metrics, through which it can define "What is the behavior of the metrics in object-oriented design". They may be described all of the behaviors like:

i. CCM (Class Coupling Metric),
ii. OXM (Operating Complexity Metric),
iii. OACM (Operating Argument Complexity Metric),
iv. ACM (Attribute Complexity Metric),
v. OCM (Operating Coupling Metric),
vi. CM (Cohesion Metric),
vii. CHM (Class Hierarchy of Method) and
viii. RM (Reuse Metric).

Metrics (i) and (iii) are very subjective in nature, Metrics (iv) and metric (vii) mostly involve the count of features; and metric (viii) is a Boolean (0 or 1) indicator metric. Therefore, all of the terminologies in object oriented language, consider as the basic components of the paradigm are objects, classes, attributes, inheritance, method, and message passing. They proposed all of that each object oriented metrics concept implies a programming behavior.

b) Morris Metrics

Morris *et al.* [27] proposed a metrics suite for the object-oriented metrics systems and they define the system in the form of the tree structure and the following are the Morris's complexity and cohesion metrics. Morris

defined the complexity of the object-oriented system in the form of the depth of the tree. Depth of the tree measures the number of the sub nodes of the tree. The more the number of sub nodes of tree the more complex the system. So, complexity of an object is equal to the depth of tree or total number of sub nodes.

c) Lorenz & Kidd Metrics

Lorenz & Kidd [19] proposed a set of metrics that can be grouped in four categories are size, inheritance, internal and external. Size oriented metrics for object oriented class may be focused on count of the metrics, operations and attributes of an individual class and average value of object-oriented software as a whole. Inheritance based metrics is totally concentrated in which operations that are reused through the class hierarchy. Metrics for the class intervals are totally oriented towards the cohesion, while the external metrics were used to examine and reuse. It divide the class based metrics into the broad categories like size, internal, external inheritance and the main metrics which are focused on the size and complexity are class size (CS), Number of operations overridden by a subclass (NOO), Number of operations added by a subclass (NOA), Specialization index (SI), Average operation size (OS), Operation complexity (OC), Average number of parameters per operation (NP).

D. Metrics for Object-Oriented Software Engineering (MOOSE) : Chidamber and Kemerer (CK) et al. [11] proposed some metrics that have generated a significant amount of interest and are currently the most well known object-oriented suite of measurements for Object-Oriented software. The CK metrics suite consists of six metrics that assess different characteristics of the object-oriented design are-

(i)Weighted Methods per Class (WMC): This measures the sum of complexity of the methods in a class. A predictor of the time and effort required to develop and maintain a class we can use the number of methods and the complexity of each method. A large number of methods in a class may have a potentially larger impact on the children of a class since the methods in the parent will be inherited by the child. Also, the complexity of the class may be calculated by the cyclomatic complexity of the methods. The high value of WMC indicates that the class is more complex as compare to the low values.

(ii)Depth of Inheritance Tree (DIT): DIT metric is used to find the length of the maximum path from the root node to the end node of the tree. The following figure shows that the value of the DIT from a simple hierarchy. DIT represents the complexity and the behavior of a class, and the complexity of design of a class and potential reuse. Thus it can be hard to understand a system with many inheritance layers. On the other hand, a large DIT value indicates that many methods might be reused. A deeper class hierarchy indicates that the more methods was used or inherited through which this making more complex to predict the behavior of the class and the deeper tree indicates that there is high complexity in the design because all of the facts contained more methods and class are involved. A deep hierarchy of the class may indicates a possibility of the reusing an inherited methods.

(iii)Number of children (NOC) - According to Chidamber and Kemerer, the Number of Children (NOC) metric may be defined for the immediate sub class coordinated by the class in the form of class hierarchy[14,15]. These points are come out as NOC is used to measure that "How many subclasses are going to inherit the methods of the parent class". The greater the number of children, the greater the potential for reuse, since inheritance is a form of reuse. The greater the number of children, the greater the likelihood of improper abstraction of the parent class. The number of children also gave an idea of the potential influence for the class which may be design.

(iv)Coupling Between Objects (CBO) – CBO is used to count the number of the class to which the specific class is coupled. The rich coupling decrease the modularity of the class making it less attractive for reusing the class and more high coupled class is more sensitive to change in other part of the design through which the maintenance is so much difficult in the coupling of classes. The coupling Between Object Classes (CBO) metric is defined as "CBO for a class is a count of the number of non-inheritance related couples with classes". It claimed that the unit of "class" used in this metric is difficult to justify, and suggested different forms of class coupling: inheritance, abstract data type and message passing which are available in object-oriented programming.

(v)Response for class (RFC) - The response set of a class (RFC) is defined as set of methods that can be executed in response and messages received a message by the object of that class. Larger value also complicated the testing and debugging of the object through which, it requires the tester to have more knowledge of the functionality. The larger RFC value takes more complex is class is a worst case scenario-value for RFC also helps the estimating the time needed for time needed for testing the class.

(vi)Lack of Cohesion in Methods (LCOM): This metric is used to count the number of disjoints methods pairs minus the number of similar method pairs used. The disjoint methods have no common instance variables in the methods, while the similar methods have at least one common instance variable. It is used to measuring the pairs of methods within a class using the same instance variable. Since cohesiveness within a class increases encapsulation it is desirable and due to lack of cohesion may imply that the class is split in to more than two or more sub classes. Low cohesion in methods increase the complexity, when it increases the

error proneness during the development is so increasing.

E. Extended Metrics For Object-Oriented Software Engineering Emoose : *W.Li* et al. *[9] proposed this metrics of the Moose model. They may be described as-*

i. **Message Pass Coupling (MPC):-** It means that the number of message that can be sent by the class operations.
ii. **Data Abstraction Coupling (DAC):-** It is used to count the number of classes which an aggregated to current class and also defined the data abstraction coupling.
iii. **Number of Methods (NOM):-** It is used to count the number of operations that are local to the class i.e. only those class operation which can give the number of methods to measure it.
iv. **Size1:-** It is used to find the number of line of code.
v. **Size2:-**It is used to count the number of local attributes & the number of operation defined in the class.

F. Metrics For Object-Oriented Design (MOOD): F.B. Abreu *et al.* [37] defined MOOD (Metrics for Object-Oriented Design) metrics. MOOD refers a structural model of the object oriented paradigm like encapsulation as (MHF, AHF), inheritance (MIF, AIF), polymorphism (POF), and message passing (COF). Each of the metrics was expressed to measure where the numerator defines the actual use of any one of the feature for a particular design [38]. In MOOD metrics model, there are two main features are methods and attributes. Attributes are used to represent the status of object in the system and methods are used to maintained or modifying several kinds of status of the objects [5].

Metrics are defined as:

(i)Method Hiding Factor (MHF): MHF is defined as the ratio of the sum of the invisibilities of all methods defined in all classes to the total number of methods defined in the system under consideration. The invisibility of a method is the percentage of the total classes from which this method is not visible.

(ii)Attribute Hiding Factor (AHF): AHF is defined as the ratio of the sum of the invisibilities of all attributes defined in all classes to the total number of attributes defined in the system under consideration.

(iii)Method Inheritance Factor (MIF): MIF is defined as the ratio of the sum of the inherited methods in all classes of the system under consideration to the total number of available methods (locally defined plus inherited) for all classes.

(iv)Attribute Inheritance Factor (AIF): AIF is defined as the ratio of the sum of inherited attributes in all classes of the system under consideration to the total number of available attributes (locally defined plus inherited) for all classes.

(v)Polymorphism Factor (PF): PF is defined as the ratio of the actual number of possible different polymorphic situation for class Ci to the maximum number of possible distinct polymorphic situations for class Ci.

(vi)Coupling Factor (CF): CF is defined as the ratio of the maximum possible number of couplings in the system to the actual number of couplings not imputable to inheritance.

MIF & AIF are used to measure the inheritance of the class & also provide the similarity into the classes. CF is used to measure the coupling between the classes. the coupling are of two types static & dynamic coupling, due to which is increase the complexity of the class & reduce the encapsulation & potential reuse that provide better maintainability. Software developers for the object-oriented system always avoid the high coupling factor. Polymorphism potential of the class are used to measure the polymorphism in the particular class & also arise from inheritance

G. Goal Question Metrics (GQM):V. L. Basili [18] developed GQM approach. This approach was originally defined for evaluating defects for a set of projects in the NASA Goddard Space Flight Center environment. He has also provided the set of sequence which are helpful for the designers. The goal of GQM is to express the meaning of the templates which covers purpose, perspective and environment; a set of guidelines also proposed for driving question and metrics. It provides a framework involving three steps:

i. List major goals of the development or maintenance project.
ii. Derive from each goal the questions that must be answered to determine if the goals are being met.
iii. Decide what must be measured in order to be able to answer the questions adequately.

Goal (Conceptual level): A goal is defined for an object, for a variety of reasons, with respect to various models of quality, from various points of view, relative to a particular environment. Objects of measurement are products, processes and resources.

Question (Operational level): A set of questions is used to characterize the way the assessment/achievement of a specific goal is going to be performed based on some characterizing model.

Metric (Quantitative level): A set of data is associated with every question in order to answer it in a quantitative way. This data can be objectives and subjective, if they depend only on the objects that can be measured and not on the viewport from which they may be taken. For example, number of versions of a document, staff hours spent on a task, size of a program.

The GQM approach define some goals, refine those goals into a set of questions, and the questions are further refined into metrics. Consider the following figure, for a particular question; G1 and G2 are two goals, Q2 in common for both of these goals. Metric M2 is required by all three questions. The main idea of GQM is that each metric identified is placed within a context, so metric M1 is collected in order to answer question Q1 to help achieve the goal G1.

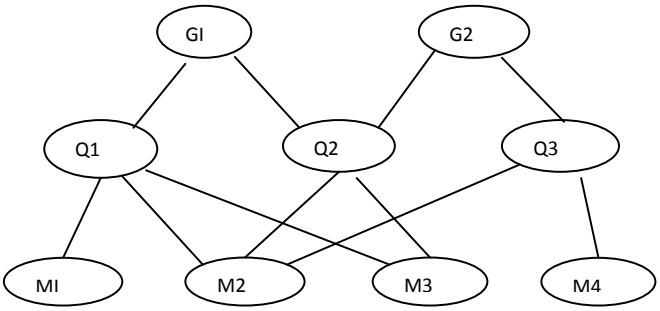


Fig 1: Goal Question Metrics Hierarchy

H. Quality Model for Object-Oriented Design (QMOOD): The QMOOD [25] is a comprehensive quality model that establishes a clearly defined and empirically validated model to assess object-oriented design quality attributes such as understandability and reusability, and relates it through mathematical formulas, with structural object-oriented design properties such as encapsulation and coupling. The QMOOD model consists of six equations that establish relationship between six object-oriented design quality attributes (reusability, flexibility, understandability, functionality, extendibility, and effectiveness) and eleven design properties.

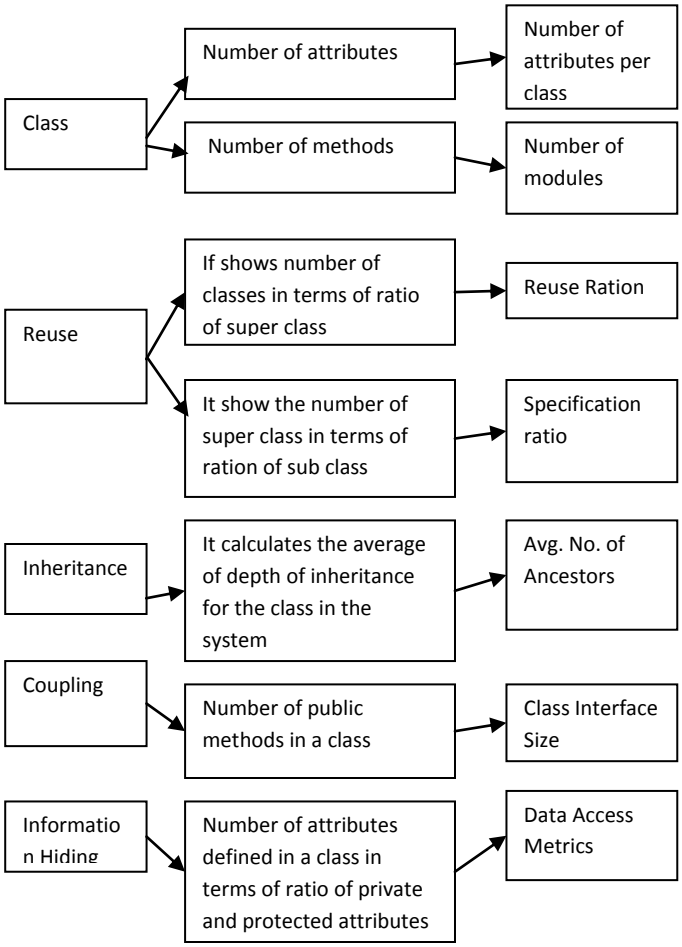


Fig 2 : QMOOD Metrics [25]

The whole description for QMOOD can be get from the Bansiya's thesis through which, The QMOOD metrics can further classified into two measures are:

System Measures: - System measures describe such metrics are DSC (Design Size in Metrics), NOH (Number of Hierarchies), NIC (Number of Independent classes), NSI (Number of Single Inheritance), NMI (Number of multiple Inheritance), NNC (Number of Internal Classes), NAC (Number of Abstract Classes), NLC (Number of Leaf Classes), ADI (Average Depth of Inheritance), AWI (Average Width of Classes), ANA (Average Number of Ancestors).

Class Measures:- Class measure metrics are those metrics which can define some metrics are MFM (Measure of Functional Modularity), MFA (Measure of Functional Abstraction), MAA (Measure of Attribute Abstraction), MAT (Measure of Abstraction), MOA (Measure of Aggregation), MOS (Measure of Association), MRM (Modeled Relationship Measure), DAM (Data Access Metrics), OAM (Operation Access Metrics), MAM (Member Access Metrics), DOI (Depth of Inheritance), NOC (Number of Children), NOA (Number of Ancestor), NOM (Number of Methods), CIS (Class

Interface Size), NOI (Number of Inline Method), NOP (Number of Polymorphic Method), NOO (Number of Overloaded Operators), NPT (Number of Unique Parameter Types), NPM (Number of Parameter per Method), NOA (Number of Attributes), NAD (Number of Abstract Data Types), NRA (Number of Reference Attributes), NPA (Number of Public Attributes), CSB (Class Size in Bytes), CSM (Class Size in Metrics), CAM (Cohesion Among Methods of class), DCC (Direct Class Coupling), MCC (Maximum Class Coupling), DAC (Direct Attribute based Coupling), MAC (Maximum Attribute based Coupling), DPC (Directed Parameter based Coupling), MPC (Maximum Parameter based Coupling), VOM (Virtual ability Of Method), CEC (Class Entropy Complexity), CCN (Class Complexity based on Data), CCP (Class Complexity based on method Parameter), CCM (Class Complexity based on Members).

IV. Li W. Metrics

Li *et al.* [16] proposed six metrics are Number of Ancestor Classes (NAC), Number of Local Methods (NLM), Class Method Complexity (CMC), Number of Descendent Classes (NDC), Coupling Through Abstract data type (CTA), and Coupling through Message Passing (CTM).

(i)Number of Ancestor Classes (NAC):- The Number of Ancestor classes (NAC) metric proposed as an alternative to the DIT metric measures the total number of ancestor classes from which a class inherits in the class inheritance hierarchy. The theoretical basis and viewpoints both are same as the DIT metric. In this the unit for the NAC metric is "class", justified that because the attribute that the NAC metric captures is the number of other classes' environments from which the class inherits.

(ii)Number of Local Methods (NLM) - The Number of Local Methods metric (NLM) is defined as the number of the local methods defined in a class which are accessible outside the class. It measures the attributes of a class that WMC metric intends to capture. The theoretical basis and viewpoints are different from the WMC metric. The theoretical basis describes the attribute of a class that the NLM metric captures. This attribute is for the usage of the class in an object-oriented design because it indicates the size of a class's local interface through which other classes can use the class. They stated three viewpoints for NLM metric as following:

1) The NLM metric is directly linked to a programmer's effort when a class is reused in an Object-Oriented design. More the local methods in a class, the more effort is required to comprehend the class behavior.
2) The larger the local interface of a class, the more effort is needed to design, implement, test, and maintain the class.
3) The larger the local interface of a class, the more influence the class has on its descendent classes.

(iii)Class Method Complexity (CMC) - The Class Method Complexity metric is defined as the summation of the internal structural complexity of all local methods. The CMC metric's theoretical basis and viewpoints are significantly different from WMC metric. The NLM and CMC metrics are fundamentally different as they capture two independent attributes of a class. These two metrics affect the effort required to design, implement, test and maintain a class.

(iv)Number of Descendent Classes (NDC) - The Number of Descendent Classes (NDC) metric as an alternative to NOC is defined as the total number of descendent classes (subclass) of a class. The stated theoretical basis and viewpoints indicate that NOC metric measures the scope of influence of the class on its sub classes because of inheritance. Li claimed that the NDC metric captures the classes attribute better than NOC.

(v)Coupling through Abstract Data Type (CTA) – The Coupling through Abstract Data Type (CTA) is defined as the total number of classes that are used as abstract data types in the data-attribute declaration of a class. Two classes are coupled when one class uses the other class as an abstract data type [16]. The theoretical view was that the CTA metric relates to the notion of class coupling through the use of abstract data types. This metric gives the scope of how many other classes' services a class needs in order to provide its own service to others.

(vi)Coupling through Message Passing (CTM) - The Coupling through Message Passing (CTM) defined as the number of different messages sent out from a class to other classes excluding the messages sent to the objects created as local objects in the local methods of the class. Two classes can be coupled because one class sends a message to an object of another class, without involving the two classes through inheritance or abstract data type [Li., 98]. Theoretical view given was that the CTM metric relates to the notion of message passing in object-oriented programming. The metric gives an indication of how many methods of other classes are needed to fulfill the class' own functionality.

a) SATC's Metrics

Rosenberg Linda [48] proposed to select object oriented metrics that supports the goal of measuring the code, quality, result and they proposed many object-oriented metrics due to lack of theoretical basis and that can be validated. These metrics may be used to evaluate the object oriented concepts like methods, coupling and inheritance and mostly focus on both of the internal and external efficiency measures of the psychological complexity factors that affect the ability of the programmer. It proposed three traditional metrics

and six new metrics for the object-oriented system metrics-

b) Traditional Metrics

(i)Cyclomatic Complexity (CC): - Cyclomatic Complexity is used to measure the complexity of an algorithm in a method of class. Cyclomatic Complexity of methods can be combined with other methods to measure the complexity of the class. Generally, this is only used for the evaluation of quality attribute complexity.

(ii)Line of Code: - It is a method used to evaluate the ease of understandability of the code by the developer and the maintainer. It can easily be counted by the counting the number of lines for the code and so on. Generally, used to measure the reusability and maintainability.

c) New Object Oriented Metrics

The six new object oriented metrics are may be discussed as:

(i)Weight Method per Class (WMC):- It is used to count the methods implemented within a class. The number of methods and complexities involved as predictors, how many time and effort is required to develop and maintain the class.

(ii)Response for a Class (RFC):- It is used to the combination of the complexity of a class through the number of methods and the communication of methods with other classes. This is used to evaluate the understandability and testability.

(iii)Lack of Cohesion of Method (LCOM):- Cohesion is a degree of methods through which all the methods of the class are inter-related with one another and provide a well bounded behavior. It also measures the degree of similarity of methods by data inputs variables and attributes. Generally, ii is used to evaluate the efficiency and reusability.

(iv)Depth of Inheritance Tree (DIT):- Inheritance is a relationship between the class that enables the programmer to use previously defined object including the operators and variables. It also helps to find out the inheritance depth of the tree from current node to the ancestor node. It is used to evaluate the reusability, efficiency, understandability and testability.

(v)Number of Children (NOC):- his is used to measure the subclass subordinate to a class in the hierarchy. Greater the number of children means greater reusability and inheritance i.e. in the form of reuse. Generally, it is used to measure efficiency, testability and reusability.

SATC focused on some selected criteria for the object oriented metrics as:

i. Efficiency of constructor design to decrease architecture complexity.
ii. Specification of design and enhancement in testing structure
iii. Increase capacity of psychological complexity.

Source Construct	Metrics	Object-Oriented Structure
Traditional metrics	Cyclomatic complexity (CC)	Methods
	Line of Codes	Methods
	Comment percentage (COM)	Methods
New object oriented metrics	Weight method per class (WMC)	Methods / Class
	Response for a class (RFC)	Class / Message
	Lack of cohesion of methods (LCOM)	Class / Cohesion
	Coupling between Object (CBO)	Coupling
	Depth of Inheritance Tree (DIT)	Inheritance
	Number of children (NOC)	Inheritance

V. Comparison Table

There is a comparison table through which, we can compare all the metrics with the multiples number of methods which are using in object-oriented design. These metrics can help for a software developer to measure the size, complexity and efforts by using these metric. They may be represented as-

Source / Metrics	CHEN	MORRIS	L & K	MOOSE	EMOOSE	MOOD	GQM	QMOOD	LIW	SATC
CCM	Y	-	-	-	-	-	-	Y	-	Y
OXM	Y	-	-	-	-	-	-	-	-	-
OACM	Y	-	-	-	-	-	-	-	-	-
ACM	Y	-	-	-	-	-	-	-	-	-
OCM	Y	-	-	-	-	-	-	-	-	-
CM	Y	-	-	-	-	-	-	-	-	-
CHM	Y	-	-	-	-	-	-	-	-	-
RM	Y	-	-	-	-	-	-	-	-	-
DIT	-	Y	-	Y	-	-	-	Y	-	Y
LCOM	-	Y	-	Y	-	-	-	-	-	Y
CBO	-	Y	-	Y	-	-	-	-	-	-
CS	-	-	Y	-	-	-	-	-	-	-
NOA	-	-	Y	-	-	-	-	Y	-	-
NOO	-	-	Y	-	-	-	-	Y	-	-
SI	-	-	Y	-	-	-	-	-	-	-
OS	-	-	Y	-	-	-	-	-	-	-
OC	-	-	Y	-	-	-	-	-	-	-
NP	-	-	Y	-	-	-	-	-	-	-
WMC	-	-	-	Y	-	-	-	-	-	Y

NOC	-	-	-	Y	-	-	-	Y	-	Y
RFC	-	-	-	Y	-	-	-	-	-	Y
MPC	-	-	-	-	Y	-	-	Y	-	-
DAC	-	-	-	-	Y	-	-	Y	-	-
NOH	-	-	-	-	Y	-	-	-	-	-
MHF	-	-	-	-	-	Y	-	-	-	-
AHF	-	-	-	-	-	Y	-	-	-	-
MIF	-	-	-	-	-	Y	-	-	-	-
AIF	-	-	-	-	-	Y	-	-	-	-
PF	-	-	-	-	-	Y	-	-	-	-
CF	-	-	-	-	-	Y	-	-	-	-
NOH	-	-	-	-	-	-	-	Y	-	-
NIC	-	-	-	-	-	-	-	Y	-	-
NMI	-	-	-	-	-	-	-	Y	-	-
NNC	-	-	-	-	-	-	-	Y	-	-
CCP	-	-	-	-	-	-	-	Y	-	-
NAC	-	-	-	-	-	-	-	-	Y	-
NLM	-	-	-	-	-	-	-	-	Y	-
CMC	-	-	-	-	-	-	-	-	Y	-
NDC	-	-	-	-	-	-	-	-	Y	-
CTA	-	-	-	-	-	-	-	-	Y	-
CTM	-	-	-	-	-	-	-	-	Y	-

VI. Conclusion and Future Works

This manuscript contributes to an increased understanding of the state of the software metrics. A mechanism is provided for comparing all the object oriented software metrics which define all the methods, attributes are used in software engineering environment. The increase is software development means the measurement was also so high. The increasing significance being placed software measurement which has to lead and increase amount of research on developing the new software measures. In this paper, we have presented all of the software metrics for object oriented development. They provided a basis for measuring all of the characteristics like size, complexity, performance and quality. In rely of some notions the quality may be increased by added some features like abstraction, polymorphism and inheritance which are inherent in object orientation. This paper provides some help for researchers and practitioners for better understanding and selection of software metrics for their purposes.

References Références Referencias

1. C. Neelamegam, M. Punithavali, "A survey on object oriented quality metrics", Global journal of computer science and technologies, pp 183-186, 2011.
2. A. Deepak, K. Pooja, T. Alpika, S. Sharma, "Software quality estimation through object oriented design metrics", IJCSNS International journal of computer science and network security, april 2011, pp 100-104.
3. A. Henderson, seller, "object oriented metrices: measure of complexity", Prentice Hall, 1996.
4. A. Shaik, C.P.K. Reddy, B. Manda, prakashine, K. Deepti, "Metrics for object oriented design software system:A Survey", Journal of emerging trend in engineer and applied science (JETEAS), pp 190-198, 2010
5. N. Fenton et al, "Software metrices: a rigorous and practical approach", International Thomson computer press 1996.
6. L.C.Briand, J.Wuest, J.Daly and V. Porter, "Exploring the Relationships Between Design Measures and Software Quality In Object Oriented Systems", Journal of Systems and Software, 51, 2000.
7. L.C. Briand, W.L. Melo and J.Wust, "Assessing the Applicability of Fault Proneness Models Across Object Oriented Software Projects", IEEE transactions on Software Engineering. Vol. 28, No. 7, 2002.
8. P.Coad and E.Yourdon, "Object Oriented Analysis", Yourdon Press, 1990.
9. W. Li, Sallie, Henry "Metrics for Object-Oriented system", Transactions on Software Engineering, 1995.
10. L.H. Rosenberg and L.Hyatt, "Applying and interpreting object oriented metrics", Proceedings of software technology conference, utah, April 1998.
11. A. Shyam, Kemerer, F. Chris, "A Metrics Suite for Object- Oriented Design" M.I.T. Sloan School of Management, pp. 53-315, 1993.
12. R. Harrison, Samaraweera, L.G. Dobie and Lewis, P.H: Comparing Programming Paradigms: An Evaluation of Functional and Object-Oriented Programs, Software Eng. J., vol. 11, pp. 247-254, July 1996.
13. M. Alshayeb and Li.W.,"An empirical validation of object-oriented metrics in two different iteration software processes", IEEE transactionod Software Engineering, Vol-29, no-.11, Nov 2003.
14. A. Shyam and C. F. Kemerer, "Towards a Metrics Suite for Object Oriented Design", Proceeding on Object Oriented Programming Systems, Languages and Applications Conference (OOPSLA'91), ACM, Vol. 26, Issue 11, Nov 1991, pp. 197-211.
15. C. Shyam and C. F. Kemerer, "A Metrics Suite for Object Oriented Design", IEEE Transactions on Software Engineering, Vol. 20, No. 6, June 1994, pp. 476-493.
16. Li W., "Another Metric Suite for Object-oriented Programming", The Journal of System and Software, Vol. 44, Issue 2, December 1998, pp. 155-162.
17. C. Jones, "Estimating Software Costs: Bringing Realism to Estimating", 2nd Edition, Mc Graw Hill, New York, 2007.
18. V.L.Basili, L. Briand and W. L. Melo, "Avalidation of object-oriented Metrics as Quality Indicators", IEEE Transaction Software Engineering. Vol. 22, No. 10, 1996, pp. 751-761.

19. M. Lorenz, J. Kidd, "Object Oriented Software Metrics", Prentice Hall, NJ, (1994).
20. R. D. Neal, "The Measurement Theory Validation of Proposed Object-Oriented Software Metrics", Dissertation, Virginia Commonwealth University, (1996).
21. V. Gurp, J. Bosch, "Design, Implementation and Evolution of Object-Oriented Frameworks: Concepts and Guidelines", Software- Practice and Experience, 31, (3), (2001), 277–300.
22. H. Fujiwara, S. Kusumoto, K. Inoue, A. Suzuki, T. Ootsubo, K. Yuura, "Case Studies To Evaluate a Domain Specific Application Framework Based on Complexity and Functionality Metrics", Information and Software Technology, 45, (1), (Jan 2003), 43–49.
23. H. Etzkorn, W. E. Hughes, C. G. Davis, "Automated Reusability Quality Analysis of OO Legacy Software", Information and Software Technology, 43, (5), (2001), 295–308.
24. G. Manduchi, C. Taliercio, "Measuring Software Evolution at a Nuclear Fusion Experiment Site: A Test Case for Applicability of OO and Reuse Metrics in Software Characterization", 44, (10), (2002), 593–600.
25. J. Bansiya, C. G. Davis, "A Hierarchical Model for Object-Oriented Design Quality Assessment", IEEE Transactions on Software Engineering, 28, (1), (2002), 4–17.
26. G. Subramanian, W. Corbin, "An Empirical Study of Certain Object-Oriented Software Metrics", Journal of Systems and Software, 59, (1), (2001), 57–63.
27. K. Morris, "Metrics for Object-oriented Software Development Environments," Masters Thesis, MIT, 1989..
28. Booch, G: Object-Oriented Analysis and Design with Applications, 2nd ed., Benjamin Cummings, 1994.
29. R. S. Pressman: —Software Engineering, Fifth edition, ISBN 0077096770.
30. Chen J.Y Lum: "A New Metrics for Object-Oriented Design."Information of Software Technology 35, 4(April 1993):232-240.
31. Abreu, B. Fernando: "The MOOD Metrics Set," Proc. ECOOP'95 Workshop on Metrics, 1995.
32. Alexander et al 2003,.Mathematical Assessment of Object-Oriented Design Quality., IEEE Transactions on Software Engineering, VOL. 29, NO. 11, November 2003.
33. H.Lilu, K.Zhou and S.Yang: "Quality metrics of OOD for Software development and Re-development", First Asia-Pacific Conference on Quality Software, August 2002.
34. M.Subramanyam and R.Krishnan: "Emphirical Analysis of CK metrics for OOD complexity: Implication for software defect", IEEE transaction on software engineering, 2003.
35. R.Harrison, S.J.Counsell and R.V.Nithi: "An evaluation of the MOOD set of OOSM", IEEE Transaction on Software Engineering, vol.24 no.6, pp.491-496, June 1998. JürgenWüst, "SD METRICS TOOL", in der Lache 17, 67308.
36. J.Eder, G.Kappel and M.Schreft, "Coupling and Cohesion in ObjectOriented Systems", Technical Report University of Klagenfurt, 1994.
37. B. F. Abreu: "Design metrics for OO software system", ECOOP'95, Quantitative Methods Workshop, 1995.
38. Abreu, B. Fernando, Rita, Miguel, G.: "The Design of Eiffel Program: Quantitative Evaluation Using the MOOD metrics", Proceeding of TOOLS'96 USA, Santa Barbara, California, July 1996.
39. A. Kaur, S. Singh and K. S. Kahlon, Evaluation and Metrication of Object Oriented System", Proceedings of the International Multi Conference of Engineers and Computer Scientists 2009, I IMECS 2009, March 18 - 20, 2009, Hong Kong.
40. F.Brito, E. Abreu and W. Melo, 1996, "Evaluating the Impact of Object-Oriented Design on Software Quality", 3rd Int'l S/W Metrics Symposium, March 1996, Berlin, Germany
41. S.R.Chidamber and C.F.Kemerer. A metrics suite for object oriented design. IEEE Transactions on Software Engieneering , pages 476 – 493, June 1994.
42. Dr B.R. Sastry, M.V. Vijaya Saradhi, "Impact of software metrics on Object Oriented Software Development life cycle", International Journal of Engineering Science and Technology, Vol 2 (2), pg 67-76, 2010.
43. R. Subramanyam and M.S. Krishnan, " Empirical Analysis of CK metricsfor Object Oriented Design Complexity: Implications of Software defects" IEEE transactions on Software Engineering, vol 29, no- 4, 2003.
44. Dr K. M. Breesam, "Metrics for Object Oriented design focusing on class Inheritance metrics", 2nd International conference on dependability of computer system IEEE, 2007.
45. C. Neelamegan, Dr. M. Punithavalli, "A Survey-Object Oriented quality metrics", Global journal of Computer Sc. And Technology, Vol 9, no 4, 2009.
46. A. Goldberg, Robson, D., "Smalltalk-80: the language and its implementation", Reading, MA: Addison Wesley, 1983.
47. M.A. Eliss & B. Stroustrup, "The annotated C++ Reference manual", Reading MA: Addison-Wesley, 1990.
48. Rosenberg Linda, "Software Quality Metrics for Object Oriented System Environments", A report of SATC's research on OO metrics.
49. A. Brito, E. Abreu (INESC/ISEG), Miguel Goulao, Rita Esteves (INESC/IST), "Towards the Design Quality Evaluation of Object Oriented Software

System", Revised version: originally published in Proceedings of 5th International Conference of Software Quality, Austin, Texas, 23 to 26 October, 1993.

50. A. L. Baroni, "Foramal Definition of Object Oriented Design Metrics", MS Thesis, Vrije Universities Brussels-belgium, 2002.
51. Abreu, Brito F. and Carapuca, Rogerio, "Candidate Metrics for Object –Oriented Software within a Taxonomy Framework", Proceddings of AQUIS'93 (Achieving Quality In Software), Venice, Italy, October 1993: selected for reprint in the Journals of Systems and Software, Vol. 23 (I), pp. 87-96, July 1994.
52. S. Chidamber and C. Kemerer, "A metrics Suite for Object Oriented Design", IEEE Transaction on Software Engineering, June 1994, pp 476-492.
53. J. Xu, H. Danny, L. Fernando Capretz, An Empirical Validation of Object-Oriented Design Metrics for Fault Prediction,. Jo] urnal of Computer Science, Vol: 4, No: 7, pp. 571-577, 2008.
54. Y. Zhou, H. Leung, "Empirical Analysis of Object-Oriented Design Metrics for Predicting High and Low Severity Faults", IEEE transaction on software engineering, vol. 32, no.10, pp-771-789,2006.

Appendix

CCM	Class coupling metrics
OXM	Operating complexity metrics
OACM	Operating argument complexity metrics
ACM	Attribute complexity metrics
OCM	Operating coupling metrics
CM	Cohesion metrics
CHM	Class hierarchy of methods
RM	Reuse metrics
DIT	Depth of inheritance tree
LCOM	Lack of cohesion in methods
CBO	Coupling between objects
CS	Class size
NOA	Number of operation added by some class
NOO	Number of operation overridden by subclass
SI	Specialization index
OS	Average operation size
OC	Operation complexity
NP	Average number of parameter per operation
WMC	Weighted method per class
NOC	Number of children
RFC	Response for class
MPC	Message pass coupling
DAC	Data abstraction coupling
NOH	Number of methods
MHF	Methods hiding factor
AHF	Attribute hiding factor
MIF	Methods inheritance factor
AIF	Attribute hiding factor
PF	Polymorphism factor
CF	Coupling factor
NOH	Number of hierarchies
NIC	Number if independent class
NMI	Number of multiple inheritance
NIC	Numbers of inheritance class
CCP	Class complexity based on methods parameter
NAC	Number of ancestor class
NLM	Number of local methods
CMC	Class methods complexity
NDC	Number of descendent class
CTA	Coupling through abstract data types
CTM	Coupling through message passing

Global Journal of Computer Science and Technology
Volume 12 Issue 7 Version 1.0 April 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

OSSM: Ordered Sequence Set Mining for Maximal Length Frequent Sequences A Hybrid Bottom-Up-Down Approach

By Anurag Choubey, Dr. Ravindra Patel & Dr. J.L. Rana
Rajiv Gandhi Technological University, Bhopal, India

Abstract - The process of finding sequential rule is indispensable in frequent sequence mining. Generally, in sequence mining algorithms, suitable methodologies like a bottom–up approach is used for creating large sequences from tiny patterns. This paper proposed on an algorithm that uses a hybrid two-way (bottom-up and top-down) approach for mining maximal length sequences. The model proposed is opting to bottom-up approach called "Concurrent Edge Prevision and Rear Edge Pruning (CEG&REP)" for itemset mining and top-down approach for maximal length sequence mining. It also explains optimality of top-to-bottom approach in deriving maximal length sequences first and lessens the scanning of the dataset.

GJCST Classification: F.3.m

Strictly as per the compliance and regulations of:

OSSM: Ordered Sequence Set Mining for Maximal Length Frequent Sequences

A Hybrid Bottom-Up-Down Approach

Anurag Choubey [α], Dr. Ravindra Patel [σ] & Dr. J.L. Rana [ρ]

Abstract - The process of finding sequential rule is indispensable in frequent sequence mining. Generally, in sequence mining algorithms, suitable methodologies like a bottom–up approach is used for creating large sequences from tiny patterns. This paper proposed on an algorithm that uses a hybrid two-way (bottom-up and top-down) approach for mining maximal length sequences. The model proposed is opting to bottom-up approach called "Concurrent Edge Prevision and Rear Edge Pruning (CEG&REP)" for itemset mining and top-down approach for maximal length sequence mining. It also explains optimality of top-to-bottom approach in deriving maximal length sequences first and lessens the scanning of the dataset.

I. Introduction

Researchers feel enthusiastic on the sequential pattern mining problems and wide range of possibilities of applications regarding the envisaging of the customer buying patterns and scientific discoveries [1, 2, 3, 4, 5] discussed by Agrawal and Srikant [1]. Let us explain with an example like finding the given time stamped sequences of purchase made by a customer. In this example the main objective is to find sequence of same time stamped list of items purchased by the customer. So the algorithm of sequence pattern mining should concentrate on finding the repeated sequences which are called as frequent sequence. Such sequences list out the frequency of common occurrences. several heuristics like GSP [1], SPADE [3], Prefix Span [2] and the SPIRIT [4] attempt to find the frequent patterns in productive method by striving to cut short a series, hence decrease search space. The GSP algorithm [1] utilizes the anti-monotone property (all subsequences of a frequent sequence are also frequent).

The SPADE finds frequent sequences using the lattice search [3] and intersection based approach. In this particular method the sequence database is converted into a vertical format. The candidate sequences will be made into different groups. These frequent sequences will be listed in SPADE utilizing two methods namely breadth first method and depth first method. The base is calculated for the produced sequences. The approach of mentioned three algorithms can be grouped as the candidate-production with a base evaluation. The PrefixSpan [2] algorithm adopts growth method pattern. It utilizes recorded database to accomplish.. Prefix is Projected Sequential Pattern mining which checks prefix subsequences and includes the postfix sub sequences into the databases.

Author α : Dean Academic, Technocrats Institute of Technology, Bhopal M.P., India. E-mail: anuragphd11@gmail.com
Author σ : Associate Professor and Head, Department of Computer Applications at Rajiv Gandhi Technological University, Bhopal, India. Email: ravindra@rgtu.net
Author ρ : Group Director, Radha Raman Group of Institute, Bhopal. Email: jl_rana@yahoo.com

II. Related Work

The sequential item set mining problem was initiated by Agrawal and Srikant, and the same developed a filtered algorithm, GSP [1], based on the Apriori property [1]. Since then, lots of sequential item set mining algorithms are being developed to increase the efficiency. Some are SPADE [3], PrefixSpan[2], and SPAM [11]. SPADE[3] is on principle of vertical id-list format and it uses a lattice-theoretic method to decompose the search space into many tiny spaces, on the other hand PrefixSpan[2] implements a horizontal format dataset representation and mines the sequential item sets with the pattern-growth paradigm: grow a prefix item set to attain longer sequential item sets on building and scanning its database. The SPADE[3] and the PrefixSPan[2] highly perform GSP[1]. SPAM[11] is a recent algorithm used for mining lengthy sequential item sets and implements a vertical bitmap representation. Its observations reveal, SPAM[11] is more efficient in mining long item sets compared to SPADE[3] and PrefixSpan[2] but, it still takes more space than SPADE[3] and PrefixSpan[2]. Since the frequent closed item set mining [12], many capable frequent closed item set mining algorithms are introduced, like A-Close [12], CLOSET [13], CHARM [14], and CLOSET+ [15]. Many such algorithms are to maintain the ready mined frequent closed item sets to attain item set closure checking. To decrease the memory usage and search space for item set closure checking, two algorithms, TFP [17] and CLOSET+2, implement a compact 2-level hash indexed result-tree structure to keep the readily mined frequent closed item set candidates. Some pruning methods and item set closure verifying methods, initiated that can be extended for optimizing the mining

of closed sequential item sets also. CloSpan is a new algorithm used for mining frequent closed sequences [16]. It goes by the candidate maintenance-and-test method: initially create a set of closed sequence candidates stored in a hash indexed result-tree structure and do post-pruning on it. It requires some pruning techniques such as Common Prefix and Backward Sub-Itemset pruning to prune the search space as CloSpan[16] requires maintaining the set of closed sequence candidates, it consumes much memory leading to heavy search space for item set closure checking when there are more frequent closed sequences. Because of which, it does not scale well the number of frequent closed sequences.

III. Sequence, Sub Sequence and Frequent Sequences

We can say a sequence means an ordered set of events [1], set of events $S=\{s_1,s_2,s_3,....s_j,s_{j+1},.....s_n \mid \exists! s_i\, i=\{1,2,3,\ j,\ j+1,....n\}\}$ And every event s_i is considered as an item set, which is a non-empty, unordered, finite set of items, which can be represented as $s_i=\{i_1,i_2,i_3,.....,i_m \mid \exists! i_e\, e=\{1,2,3,4,....,m\}\}$, here $i_e\ for\ e=1..m$ is item. The length of a sequence is the number of items present in the sequence. A sequence can be referred with its length, as an example a sequence of length k is called a k-sequence. A sequence $S1$ is said to be a subsequence of another sequence $S2$ if and only if $s1_i \subseteq s2_{e_i}$ for $\mathrm{i}=\{1,2,3,\ldots.\mathrm{m} \mid e_1<e_2<e_3.....<e_i.....<i_n\}$ and e_i is event of $S2$ .The $S2$ is said to super sequence of $S1$ and $S1$ is said to subsequence of $S2$. The sequence database S is a set of the form (tid, s) where tid is the transaction-id and s is the sequence generated from transaction.

Let the minimum support as a threshold defined by user which indicates the desired minimum occurrences of the sequence to be claimed as frequent.

A sequence s_j is lengthiest if $s_j! \subseteq \{s_i \in S \mid i=\{1..m\}\}$. We explain the maximal support of an event e which consists of items $(i_1,i_2,...,i_k)$ as

$$MS_{e_i}=\max\{\sup(i_1),\sup(i_2),\sup(i_3),.....,\sup(i_m) \mid \exists! i_t\ t=\{1,2,3,....,m\}\}$$

Here $\sup(i_n)$ frequency of item i_n occurrence. And, the maximal support factor MSF of a sequence is represented by sum of maximal support of all events belongs to that sequence. Let s be a sequence of events $\{e_1,e_2,.....,e_m \mid \exists! e_i\ i=\{1,2,3,....,m\}\}$ and then maximal support factor can be measured as

$$MSF(s)=\sum_{i=1}^{m} MS(e_i)$$

The Maximal Support Factor is the threshold used in our proposal to minimize the subsequence search in top to bottom approach.

We apply an ordered search on sequence database, hence the sequence database will be ordered in descending manner by MSF of sequences. Then the search for super sequence of a given sequence is limited to the sequences with greater or equal MSF of the given sequence.

The preprocessed dataset with the sorted transaction list has several properties that can be used to cut short the search space which are hypothesized below:

- **Hypothesis 1:** A super sequence search of given sequence s_c can stop at a sequence s_t if $MSF(s_t)<MSF(s_c)$. Putting differently a candidate sub sequence s_c will not be a subsequence of s_i if $MSF(s_c)>MSF(s_i)$.
- **Hypothesis 2:** A sequence s_c is frequent if count of super sequences for s_c is equal or greater than the given minimum support threshold by user. $\mathtt{minsupport(s_C)}$ can be referred as $\mathrm{maxOccurrence(s_c)}$ and it can be measured as

$$\max Occurrence(s_C)=\sum_{i=1}^{|SM|}\{i \mid if\ (s_C \subseteq s_i)\, i=1 or\ i=0\}$$

- Here $|SM|$ is set of order sequences in descending manner for each sequence MSF is greater or equal to $MSF(s_c)$.
- **Hypothesis 3:** Let s_m be maximal sequence and $\mathrm{maxOccurrence}(s_m) \geq ms$ then avoid all subsequences of s_m from considering to evaluate maximal sequences
- **Hypothesis 4:** Let s_c be sequence such that $\mathrm{maxOccurrence}(s_c)<ms$ then discard all supersets of s_c .Since s_c is infrequent then its super sequences also infrequent.

IV. Ossm Overview

- Initially we apply (CEG & REP)[7] to find closed frequent itemsets, which is bottom-up approach called Concurrent Edge Prevision and Rear Edge Pruning (CEG & REP)[7].
- Eliminate the items from transactions that are not part of any event e such that $e \in I$ and referred as outliers. An item i is an outlier if
  - $i \in T_{tid}$ and $i \notin \{I_s \mid for\ s = \{1.....\mid I \mid\}$
  - Here T_{tid} is a transaction represented by transaction id tid and I_s is an itemset that belongs to set I of frequent itemsets.
- Build sequences by grouping items as events in a given transaction. Here we follow the top-to-bottom approach to build events. First we build events based on maximal length itemsets and continue the process in descending order of the itemset lengths.
- Measure the Maximal support of each event e of the given transaction T_{tid} (refer section 3 for measuring maximal support MS for an event e of transaction T_{tid}).
- Measure the Maximal Support Factor MSF of each sequence S_{tid} (Refer section 3 for measuring Maximal Support Factor MSF of sequence S_{tid})
- Order the sequences in sequence dataset S in descending manner of their MSF.
- Build weighted acyclic directed graph from frequent itemsets of length two.
  - Elements of the itemset considered as vertices
  - Support of that itemset considered as edge weight
- Apply WFI algorithm to find critical path between any two items which represents the maximal sequence between two elements opted.
- Apply all four properties that are hypothesized in section 3 to discard, prune sequences or select maximal length sequences.

V. Ossm: Top-Down Ordered Sequence Set Mining For Maximal Length Sequences

a) *Concurrent Edge Prevision and Rear Edge Pruning (CEG&REP)[7] in OSSM*

i. *Preprocess*

Dataset preprocessing and itemsets Database initialization is performed by us as the first stage of proposal. As we find itemsets with single element, we in parallel prune it with the itemsets of single elements if the support of the selected itemsets is less than the required support.

ii. *Concurrent Edge Prevision*

In this phase, we select all itemsets from given itemset database as input in parallel. Then we start projecting edges from each selected itemset to all possible elements. The first iteration includes the pruning process in parallel, from the second iteration onwards this pruning is not required, which we claimed as an efficient process compared to other similar techniques like BIDE [8]. In first iteration, we project an itemset s_p that spawned from selected itemset s_i from DB_S and an element e_i considered from 'I'. If the $f_{ts}(s_p)$ is greater or equal to rs, then an edge will be defined between s_i and e_i. If $f_{ts}(s_i) \cong f_{ts}(s_p)$ then we prune s_i from DB_S. This pruning process is required and limited to first iteration only. From the second iteration onwards project the itemset S_p that spawned from S_p, to each element e_i of 'I'. An edge can be defined between $S_{p'}$ and e_i if $f_{ts}(s_p)$ is greater or equal to rs. In this description $S_{p'}$ is a projected itemset in previous iteration and eligible as a sequence. Then apply the following validation to find the closed sequence.

iii. *Rear Edge Pruning*

If any of $f_{ts}(s_{p'}) \cong f_{ts}(s_p)$ that edge will be pruned and all disjoint graphs except s_p will be considered as closed sequence and moves it into DB_S and remove all disjoint graphs from memory.

The termination of above process do not take place till the graph becomes empty, i.e. till the elements which are connected through transitive edges and projecting itemsets are available in the memory.

b) *Building Sequence Set from Transaction Dataset*

Here in this section we explore the process of building sequence dataset.

TD is the given transaction dataset

I is the set of closed frequent itemsets of length $1\ to\ m$. Here m is the maximal length of the itemset.

- Initially set of closed frequent itemsets is ordered in descending manner by itemset length.
- For each transaction T_{tid} in the given transaction dataset TD
  - Build events based on the closed frequent itemsets of the set I such that the event lengths will be decided in the descending order of the frequent itemset length.
  - Initially events with length of m determined that is maximal length of the frequent itemsets. Then

events with length of $m-1$ will be determined. This process continues to determine events with length $\{m-i \mid i=\{2,3,....m-1\}\}$.

- Then eliminates the items that are not part of the any event in the transaction T_{tid}, which also referred as outliers.

c) Measuring MSF and Ordering the Sequence Dataset

As a part of the OSSM, we order the sequence dataset in the descending manner of the Maximal support Factor (refer section 3 for details).

Let S be the determined sequence dataset from the given transaction dataset TD and set of frequent itemsets I.

- For each sequence s of the given sequence set S
  - Find Maximal support $MS(e)$ of the each event e of the sequence s (refer section 3 for process of measuring $MS(e)$)
  - Find Maximal Support Factor $MSF(s)$ (refer section 3 for process of measuring $MSF(s)$)
- Order the sequence set in descending order of the MSF

d) Building Weighted Acyclic Directed Graph

In this phase of OSSM, we explore the building of a weighted acyclic directed graph.

Let I_2 be the set of frequent itemsets of length 2.

Let G be the graph initially with vertex count of zero $|V|=0$ and edge count of zero $|E|=0$. Here V is vertex set and E is edge set.

- For each itemset $i_{s\to d}$ of I_2 build an edge ed in graph G
  - Let consider item $s \in i_{s\to d}$ as source vertex, and add item s to vertex set V. Increment $|V|$ by 1.
  - Let consider item $d \in i_{s\to d}$ as destination vertex, and add item d to vertex set V. Increment $|V|$ by 1.
  - Build a directed edge ed between s and d, add directed edge ed to edge set E. Increment $|E|$ by 1.
  - Add support of $i_{s\to d}$ as weight to edge ed.

e) Finding Critical Paths between two Items as Maximal Length Sequences

In this phase we apply WFI algorithm [9, 10]. In the first pass of algorithm we try to identify and evaluate potential long and rich candidates. The rich sequences are the one whose constituent 2-sequences have high support. In the directed graph, the 2-sequenece frequencies are represented by the edge weights; we can easily compute the path with the highest weights between all pairs of nodes. Here we use WFI algorithm [9,10] for the purpose finding critical path(a path with maximal vertex count) with maximal weights.

f) Sequence Evolution

Each critical path generated from the graph G will be considered as candidate sequence and stored in candidate sequence set css

g) Verifying Frequency of Candidate Sequence

- For each candidate sequence cs of candidate sequence set css
  - Find $MSF(cs)$
  - For each sequence s such that $\{s \in S \mid MSF(s) \geq MSF(cs)\}$ (refer hypothesis 1in section 3)
    - If $cs \subseteq s$ then increment $\sup(sc)$ by 1
  - If $\sup(cs) \geq st$ (refer the hypothesis 2 in section 3) then
    - Move cs to frequent sequence set fss
    - For each candidate sequence cs' of candidate sequence set css such that $cs' \neq cs$
    - If $cs' \subseteq cs$ then consider cs' as frequent (refer hypothesis 3 in section 3) and move cs' from css to fss
  - Else
    - For each candidate sequence cs' of candidate sequence set css such that $cs' \neq cs$
    - If $cs \subseteq cs'$ then consider cs' as not frequent and prune it from css (refer hypothesis 4 in section 3)

h) Finding Maximal Length Sequences

Let fss be the frequent sequence set generated in previous phase

- Order the fss in descending manner by Maximal support Factor MSF of the frequent sequences of fss.
- Let a Boolean factor sts as true.
- For each frequent sequence fs of the fss and sts is true
  - For each frequent sequence fs' such that $\{fs' \in fss \mid MSF(fs') \geq MSF(fs)\}$ and $fs \neq fs'$ and sts is true
    - If $fs \subseteq fs'$ then set sts as false
  - If sts is true then move frequent sequence fs to maximal length frequent sequence set $mlfss$

Finally, maximal length frequent sequence set $mlfss$ contains sequences that are not subsequence of any other frequent sequences.

VI. Conclusion

The proposed ordered sequence set mining (OSSM) approach is a scalable and optimal because of its hybrid bottom-up-down approach. OSSM is supported by our earlier Concurrent Edge Prevision and Rear Edge Pruning (CEG&REP)[7] for frequent closed itemset mining, which was proven as efficient in memory usage and scalable on dense datasets. A novel mechanism of sequence dataset generation from transaction dataset is introduced in this paper. The proposed OSSM is capable to generate the longest candidate sequence by weighted acyclic directed graph construction and also efficient and scalable to find frequent sequence set and maximal length frequent sequence set due to top-down approach. To compute the support for a candidate sequence, it uses the maximal support factor of sequences. And the order approach that ordering the sequence dataset in descending order of MSF ensures that whole data set is not scanned. Also, if the data set contains long regular sequences, it is identified early enough and thus all the subsequences of this is also marked regular and need not be evaluated. The longest possible sequence is build up by bottom up algorithms starting from 2-sequence.

References Références Referencias

1. Agrawal, R., and Srikant, R., Mining Sequential Patterns: Generalizations and Performance Improvements Proceedings of the 5th International Conference on Extending Database Technology: Advances in Database Technology p. 3 – 17, (1996).
2. Pei J, Han J. et al: "PrefixSpan: Mining Sequential Patterns Efficiently by Prefix- Projected Pattern Growth" in Int'l Conf Data Engineering,p 215-226 (2001)
3. Zaki, M. J., SPADE: An Efficient Algorithm for Mining Frequent Sequences, Machine Learning, v.42 n.1-2, p.31-60, January-February 2001.
4. Garofalakis M, Rastogi R and Shim, K, "Mining Sequential Patterns with Regular Expression Constraints", in IEEE Transactions on Knowledge and Data Engineering, vol. 14, nr. 3, pp. 530-552, (2002).
5. Antunes, C and Oliveira, A.L: "Generalization of Pattern-Growth Methods for Sequential Pattern Mining with Gap Constraints" in Int'l Conf Machine Learning and Data Mining, (2003) 239- 251
6. Lin, D I, and Kedem, Z M, "Pincer Search: A new algorithm for discovering the maximum frequent set", in International conference on Extending database technology, 1998.
7. Anurag Choubey, Dr. Ravindra Patel and Dr. J.L. Rana, "Concurrent Edge Prevision and Rear Edge Pruning Approach for Frequent Closed Itemset Mining"; (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No. 11, 2011
8. Jianyong Wang, Jiawei Han: BIDE: Efficient Mining of Frequent Closed Sequences. ICDE 2004: 79-90
9. Floyd, R.: Algorithm 97: Shortest path. Communications of the ACM 5 (1962)
10. Warshall, S.: A theorem on boolean matrices. Journal of the ACM 9 (1962)
11. M. Zaki, SPADE: An Efficient Algorithm for Mining Frequent Sequences. Machine Learning, 42:31-60, Kluwer Academic Pulishers, 2001.
12. N. Pasquier, Y. Bastide, R. Taouil and L. Lakhal, Discoving frequent closed itemsets for association rules. In ICDT'99, Jerusalem, Israel, Jan. 1999.
13. M. Zaki, and C. Hsiao, CHARM: An efficient algorithm for closed itemset mining. In SDM'02, Arlington, VA, April 2002.
14. R. Kohavi, C. Brodley, B. Frasca, L.Mason, and Z. Zheng, KDD-cup 2000 organizers' report: Peeling the Onion. SIGKDD Explorations, 2, 2000.
15. J. Han, J. Wang, Y. Lu, and P. Tzvetkov, Mining Top- K Frequent Closed Patterns without Minimum Support. In ICDM'02, Maebashi, Japan, Dec. 2002.
16. P. Aloy, E. Querol, F.X. Aviles and M.J.E. Sternberg, Automated Structure-based Prediction of Functional Sites in Proteins: Applications to Assessing the Validity of Inheriting Protein Function From Homology in Genome Annotation and to Protein Docking. Journal of Molecular Biology, 311, 2002.
17. J. Pei, J. Han, and W. Wang, Constraint-based sequential pattern mining in large databases. In CIKM'02, McLean, VA, Nov. 2002.

This page is intentionally left blank

Global Journal of Computer Science and Technology
Volume 12 Issue 7 Version 1.0 April 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Empirical Analysis for Recognition of Facial Expression the State of the Art

By Geeranjali Sharma & H K Sawant
Bharati Vidyapeeth Deemed University College Of Engineering, Pune

Abstract - facial expression recognition is limited to six basic expression and several combination.The expression are classified into emotion categories rather than another technique..It is difficult task to show all facial expressions because in everyday life six basic expression occur so frequently.Emotion is often communicated by small changes in one or two facial features,on the other hand the same facial expression may occured in more than one emotions.The presence or absence of one or more facial actions value may change its prediction.One man can show their facial expression in different manner than expressing the same facial expression by other person.The facial features value changes person to person for the same facial expression.

GJCST Classification: *I.4.8*

Strictly as per the compliance and regulations of:

Empirical Analysis for Recognition of Facial Expression the State of the Art

Geeranjali Sharma [α] & H K Sawant [α]

Abstract - facial expression recognition is limited to six basic expression and several combination.The expression are classified into emotion categories rather than another technique..It is difficult task to show all facial expressions because in everyday life six basic expression occur so frequently.Emotion is often communicated by small changes in one or two facial features,on the other hand the same facial expression may occured in more than one emotions.The presence or absence of one or more facial actions value may change its prediction.One man can show their facial expression in different manner than expressing the same facial expression by other person.The facial features value changes person to person for the same facial expression.

I. Introduction

Human computer interaction is an upcoming valuable field of science and engineering that provides such a platform or environment so that a machine can interact with human in natural ways. Computer will interact with human only if they will have communication skills like human. Emotion is pioneer communication skills. The very authentic way of expressing emotional state is through facial expression. Emotions are shown by some visual, vocal and other physiological means. The research on recognizing emotion through facial expression was pioneered by Ekman and Friesen [1]. Facial expression plays vital role to show the emotion and that be unique component to deliver the communication skills. Facial expression analysis has wide range of application in area such as Security and surveillance system, Gaming, Human behavior recognition, intelligent human computer interfacing, user behavior analysis. six type of facial expression happy(1), angry(2), surprise(3), disgust(4), fear(5), sad(6). The face expression recognition Problem is tough because different person display the same expression in different manner. Selecting the most important feature and ignoring unimportant feature is a main concept to solve this problem. In Automatic Facial Expression Recognition System face is first detected and localized than facial feature is extracted from the detected Face Region. finally the facial expression are classified based on the extracted feature. Sometimes the facial expression analysis has been confused with emotion analysis in the computer vision domain. For emotion analysis higher knowledge is required. Darwin (1872) is the first to point out the importance of facial expressions as powerful and immediate means for human beings to communicate their emotions and Tomkins (1963) reaches the same conclusion and reports that affect (or emotion) is always represented by visual sign. Though facial expression can convey emotion and can also express intention, cognitive processes, physical effort or other interpersonal meanings. Human face is the richest source of nonverbal communication and the most accessible interface displaying human emotions. To recognize the facial expressions using computer is a revolutionary work which rely on human-computer interaction so that computer will be able to understand whether users feel excited or bored, agrees or disagrees.

It will be a great challenge and a practical significance to develop a computer vision system which can recognize a variety of facial expressions and estimate expression intensity.

Facial expression analysis basically depends on the facial feature like eye, nose, mouth, eyebrow which play an important role to detect human emotions because these positions are changed when a human expression changes. So these are the important features to display the emotions .So these are the key point in facial expression analysis. The face model features are the feature used to represent (model) the face. The face can be represented in various way e.g., as a whole unit (holistic representation), as a set of features (analytic representation) or as a combination of these (hybrid approach).

II. Literature Review

In this work the facial expression is based on the attributes of facial muscle that is hidden state of a HMM for individual image. Probability of the state is changed on the behalf of the feature vector obtained from image processing.[2]

Optical flow algorithm is used for the evaluation of velocity vector of two successive frames. After that FFT is applied to a velocity vector around the region of mouth and eye. The selection of feature vector is lower frequencies value and a mixture density is applied to scaled the output probability of HMM to detect the variation in human facial expression. Mixture density is very useful to enhance the accuracy as the mixture increases. Facial expression recognition is done by using the Support Vector Machine by modifying Kernels.[1] Facial expression is also recognized with the

Author α : Department Of Information Technology, Bharati Vidyapeeth Deemed University College Of Engineering, Pune-46.

help of brain activity that is governed by the EEG signal.[3]

In this there is a proposed system that analyzes EEG system and classify them into 5 classes on the two emotional dimensions name like valance and arousal. However after using the 3-fold cross validation method the recognition rate for the valence dimension is 32% and for the 37% for the arousal and overall rate is approx 71%.

Facial expression recognition is also detected with the help Markov random Field.[4] In this the main concept is the essential factor that helpful for the expression detection is Eye and mouth expression. In this the first step is done image segmentation and skin detection for the Markov random field .second step is done for the eye and mouth feature extraction. The set of different color image is used as a training set.HLV color space that is responsible for the detection of the eyes and mouth region.3rd step module is in accordance with the detection of emotions in images with the help of edge detection and measurement of gradient of eye and mouth region.

There is also a facial expression detection is done by the using of coded form that consists of multiorientation ,multiresolution of the Gabor filters in which region of surface space and precise description of the place or region are mentions and that is aligned along the human face.[12] In this paper the similarity space is matched with the result obtained from the applying the Gabor filter and the result that is semantic value that obtained on the human observations. There is concepts that terminology is known as rank correlation which emphasize the semantic similarity and facial expression image similarity that is obtained after the applying the Gabor coding. The facial expression classification is also achieved of the frontal image by using of Eigen face [21].

In this paper whole face is not taken in consideration rather than classify the face in the regions that is beneficial for the facial expression classification and projecting that faces with the Eigen face and try to train with different types of facial expressions.Than taking the average of all different region faces showing different facial expression and after that making a mask. The important thing is that this masks fading the miss region and try to highlight the region that are changing during the different facial expression. For the recognition of the facial expression The new technique active appearance model (AAM) is used to trained the faces that is available in database that is used to represent the shape and texture variation that plays an important role in facial expression recognition. The features are those that are extracted from the parameter obtained from the AMM and is used to discriminate among the classification of different expression. The feature extraction with the help of AMM better than a simple classifier like Euclidean distance. The AMM makes a efficient method for the texture and shape to model such that It plays an important role and it is thoughtful like (SVM) support vector machine.

Facial expression has also detected with the accuracy of 85% with the help of facial feature vectors obtained with the help of Gabor filter and that feature value is convolved with the Log Gabor filter.[1]

In this whole face is taken into consideration and than its accuracy is tested over the classification with PCA principle component analysis and LDA liner discriminate analysis and the result are quite good .The result is achieved on low reso0lution image without specifying the fiducially points. Facial expression recognition is also achieved with the histogram sequence of the Local Gabor binary pattern.[17] Firstly the face image is convolved with multi orientation with the help of Gabor filter than after that Gabor coefficients map are extracted. Than after local binary pattern is applied on GCM to obtain the local Gabor binary pattern. Finally the SVM is used for the classification and the result of recognition rate is quite appreciable.

III. Facial Emotion Recognition System

Fig1 : block diagram

a) Preprocessing

In the preprocessing step the environmental and other variations that are present in different images are minimized. The operation that is performed is contrast adjustment, image scaling, image brightness and other image enhancement method has done. Sometimes noise is associated with images with the variation in signaling and pixel variation so removal of this factor has become essential to achieving the better result.

b) Segmentation

As we know that in many images processing the input is image and output is image. But for the facial expression analysis there is need of feature that is extracted from the input image. In other words we can say that input should be an image but output should be the feature value that has obtained from that input image.Segmentaion plays a major role in that direction.

Segmentation basically divides an image into its regions or objects. The subdivision depends on the level of problem solving. That is the Segmentation will stop when the the objects or the regions of interest in an application has been detected. The successes or failure of the any image processing or computer vision process depends on the Segmentations accuracy. So there is proper concentration is taken place in the finding the

probability of accurate Segmentation. Most of the Segmentation algorithm basically depends on the two main properties of intensities values.

a) Discontinuity
b) Similarity.

The first category of Segmentation is taking place when the abrupt changes in the intensity such as Edge is found in an image. so In this case image is partitioned into regions.

In the second case of Segmentation an Image is partitioned into the regions that are similar for particular predefined criteria.

In my approach a facial image is segmented into 4 region means 2×2 grid of 128×128 size.

256×256

Fig 2 : Before Segmentation

128×128

Fig 3 : After Segmentation

c) Feature Extraction

Is the important method to define the any recognition System. Without extracting the key feature point it is not feasible to define the facial expression recognition System. In feature extraction method I have applied the Gabor filter on the four segmented region of the face. There is 68×68 features are coming outside from the one face region but due to application of absolute and scaling method we evaluate the major four prominent features like Gabor real part, imaginary part, magnitude and angle are evaluated. These extracted feature value is in the range of 0 and 1.so from the one face we calculates the 16 features that is defined by the Gabor filter outcome. We use the frontal face for the facial expression analysis. There is need to define such an unique parameter that can be differentiate between different facial expression of expressers. The set of parameter that can be extracted is known as feature vector and the information achieving from the feature vector is defines an uniqueness as aspect with the extraction technique. If the feature value extracted from the one expression matches with the feature extracted with the expressions of other faces than it is not known as a good feature extraction technique and it is known as feature overlap. The feature extraction should be different in comparison with the other so there is made a correlation and on behalf of this next procedure be implemented. so there is several method adopted for feature extraction and among these the Gabor filter bank based method is good. So we can say that Gabor based feature extraction technique is excellent for facial expression analysis and avoid the feature overlap condition.

d) Support Vector Machine

SVM is a useful method for the data classification. It is easier than using the Neutral Networks.SVM provides the accuracy and fast result for the data to be classified and belongs to the particular class. In SVM the data is partitioned into the two parts that one is called Training set and other is known as the testing set and each having the instances of the attributes. Each instances having one target means class labels and several attribute. The goal of the SVM is that to produce the model which predicts the target value of instances in the testing set which are given by only attributes value. It is based on the supervised learning methods. Support vector machine has a unique property that it creates the hyperplane or a set of hyperplane which has taken into the consideration for the classification. It defines the functional margin means a good separation is achieved by the hyperplane that has the largest distance to the nearest data points of any class. Greater the margin it is easy to find better accuracy and minimal errors has found.

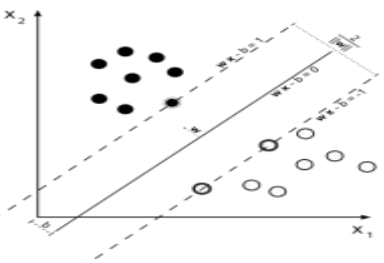


Fig 4: Hyperplane that discriminates between two classes

Lets us assume that we have Data sets of having n features of the form

$$\mathcal{D} = \{(\mathbf{x}_i, c_i) \mid \mathbf{x}_i \in \mathbb{R}^p,\ c_i \in \{-1, 1\}\}_{i=1}^{n}$$

Where C_i belongs to either -1 and +1 value indicates the which class the point X_i belongs.the equation of the hyper plane is given by in such a way

$$W \cdot X - b = 0$$

There is two cases to defining hyperplane

$$W \cdot X - b \geq 1\text{------}$$

it indicates for class one

$$W \cdot X - b \leq -1\text{-----}$$

it indicates for the class others.

Here the training vectors X_i are projected into the higher dimensional space by the function ø. the main important characteristic property of SVM is that it works on the linear separable hyperplane with the minimal margin in the high dimensional space.

C>0 is the main term to define the error parameter. The kernel function is defined as $K(\mathbf{x}_i, \mathbf{x}_j) \equiv \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j)$.

There is four basic kernels that is defined as follows:-

linear: $K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j$.

polynomial: $K(\mathbf{x}_i, \mathbf{x}_j) = (\gamma \mathbf{x}_i^T \mathbf{x}_j + r)^d,\ \gamma > 0$.

radial basis function (RBF): $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2),\ \gamma > 0$.

sigmoid: $K(\mathbf{x}_i, \mathbf{x}_j) = \tanh(\gamma \mathbf{x}_i^T \mathbf{x}_j + r)$.

Where γ, r.and d are kernel parameters.

IV. Results

Facial Expression or Facial emotion detection through image database has been proposed: Basically Image Database is Japanese Female Faces and assumed to be frontal and preprocessed and some Indian database also has been used to evaluated the accuracy, After applying Gabor Filter on the four segmented facial region we extract the Features value that is fully defined by Gabor response, The main features that are extracted known as Gabor Real part, Imaginary part, Magnitude and lastly Phase .Basically all these features value are coming out from the segmented Face region under the response of Gabor Filter, After the feature extraction maximum of absolute of Gabor filter responses on the faces are evaluated, After the feature extraction rescaling is done so that all features value are lie between the range of 0 and 1, When the features of all the classes are extracted than finally classifier Support Vector Machine (SVM) applied for labels classification, Extracted Features are arranged into the SVM format means all the value should be written in a single row and showing the name of their respective classes, Than features value of different classes is divided into two domain one for the Training set and the second one for the Testing And in last the accuracy is calculated, accuracy is above 80%.

V. Conclusion and Future Perspective

In this dissertation, we have made a facial expression recognition System that can detects the Facial Expressions like Happy, Sad, Angry, Disgust, Surprise and Fear. We have used a still images that is frontal and preprocessed. The Facial Expression analysis plays a key role in the human Machine interaction. Our approach is different to previous work done and my accuracy is above 81%.Due to segmentation and rescaling the Gabor filter output response is remarkably good. So the extracted feature detection is lie in the range of [0,+1].

This range is essential when we are going to use SVM for the classification. After using the SVM the results obtained is satisfactory. In our future perspective we will try to enhance the accuracy up to 100% and try to make a robust system that can be applied to any database whatever the alignment of the Face. In future we will develop my work for the real time system that can be utilized in any sensitive area and surveillance system. Also the Facial expression analysis is useful for the security, Gaming, Intelligent tutoring system and human behavior recognition.

References Références Referencias

1. Cohn, T. Kanade Cohn-Kanade AU-Coded Facial Expression Database Carnegie Mellon University
2. C.K. Chow, C.N. Liu. Approximating discrete probability distributions with dependence trees. IEEE Trans. Information Theory, 14:462–467, 1968.
3. I. Cohen, N. Sebe, A. Garg, L. Chen, and T.S. Huang. Facial expression recognition from video sequences: Temporal and static modeling. Computer Vision and Image Understanding, 91(1-2):160–187, 2003.
4. P. Ekman Strong evidence for universals in facial expressions. Psychol. Bull., 115(2): 268–287, 1994.
5. J.H. Friedman On bias, variance 0/1-loss, and the curse-of-dimensionality. Data Mining Knowledge Discovery, 1 (1): 55–77, 1997.
6. N. Friedman, D. Geiger, M. Goldszmidt. Bayesian network classifiers. Machine Learning, 29(2):131–163, 1997. [7] A. Garg, D. Roth. Understanding probabilistic classifiers. Proc. Eur. Conf. on Machine Learning, 179–191, 2001.
7. D. Goleman. Emotional Intelligence. Bantam Books, New York, 1995.
8. Intel Research Laboratories. OpenCV: Open computer vision library. http://sf.net/projects/opencvlibrary/.

9. C.E. Izard. Innate and universal facial expressions: evidence from developmental and crosscultural research. Psychol. Bull., 115(2): 288– 299, 1994.
10. R. Lienhart, J. Maydt. An extended set of haarlike features for rapid object detection. Proceedings of the IEEE International Conference on Image Processing, Rochester, New York, vol. 1, pp. 900-903, 2002.
11. R. Lienhart, A. Kuranov, V. Pisarevsky. Empirical Analysis of Detection cascade of Boosted Classifiers for Rapid Object Detection. Intel Corporation, Technical report, 297–304, 2002.
12. M. Pantic, L.J.M. Rothkrantz. Automatic analysis of facial expressions: the state of the art. IEEE Trans. PAMI, 22(12): 1424–1445, 2000.
13. C. Papageorgiu, M. Oren, T. Poggio. A general framework for Object Detection. Proceedings of the International Conference on Computer Vision, Bombay, India, pp. 555-562, 1998.
14. R. Schapire, Y. Freund. Experiments with a new boosting algorithm. Proceedings of the International Conference on Machine Learning, Bari, Italy, Morgan Kaufmann, pp. 148-156, 1996.
15. R. Schapire. The strenght of weak learnability. Machine Learning, 5(1), 197-227, 1990.
16. N. Sebe, I. Cohen, A. Garg, M.S. Lew, T.S. Huang. Emotion Recognition Using a Cauchy Naive Bayes Classifier. International Conference on Pattern Recognition (ICPR02), vol I, pp. 17– 20, Quebec, Canada, 2002.
17. H. Tao, T.S. Huang. Connected vibrations: a modal analysis approach to non-rigid motion tracking. Proc. IEEE Conf. on CVPR, 735–740, 1998.
18. P. Viola, M. Jones. Rapid Object Detection Using a Boosted Cascade of Simple Features. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Kauai, Hawaii, vol. 1, pp. 511-518, 2001.

This page is intentionally left blank

Global Journal of Computer Science and Technology
Volume 12 Issue 7 Version 1.0 April 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Automated Cloud Patch Segmentation of FY-2C Image Using Artificial Neural Network and Seeded Region Growing Method (ANN-SRG)

By Yu Liu, Du-Gang Xi, Xue-Gong Liu, Chun-xiang Shi & Kai Zhang
Tianjin Hydraulic Research Institute

Abstract - This paper presents a new algorithm Artificial Neural Network and Seeded Region Growing (ANN-SRG) to segment cloud patches of different types. This method used Seeded Region Growing (SRG) as segmentation algorithm, and Artificial Neural Network (ANN) Cloud classification as preprocessing algorithm. It can be trained to respond favorably to cloud types of interest, and SRG method is no longer sensitive to the seeds selection and growing rule. To illustrate the performance of this technique, this paper applied it on Chinese first operational geostationary meteorological satellite FengYun-2C (FY-2C) in three infrared channels (IR1, 10.3- 11.3μm; IR2, 11.5-12.5μm and WV 6.3-7.6μm) with 2864 samples collected by meteorologists in June, July, and August in 2007. The result shows that this method can distinguish and segment cloud patches of different types, and improves the traditional SRG algorithm by reducing the uncertainty of seeds extraction and regional growth.

Keywords : *FY-2C; multi-channel satellite image; Artificial Neural Network; Seeded Region Growing; cloud patch; segmentation; cloud type.*

GJCST Classification: *I.2.1*

Strictly as per the compliance and regulations of:

Automated Cloud Patch Segmentation of FY-2C Image Using Artificial Neural Network and Seeded Region Growing Method (ANN-SRG)

Yu Liu [α], Du-Gang Xi [σ], Xue-Gong Liu [ρ], Chun-xiang Shi [ω] & Kai Zhang [¥]

Abstract - This paper presents a new algorithm Artificial Neural Network and Seeded Region Growing (ANN-SRG) to segment cloud patches of different types. This method used Seeded Region Growing (SRG) as segmentation algorithm, and Artificial Neural Network (ANN) Cloud classification as pre-processing algorithm. It can be trained to respond favorably to cloud types of interest, and SRG method is no longer sensitive to the seeds selection and growing rule. To illustrate the performance of this technique, this paper applied it on Chinese first operational geostationary meteorological satellite FengYun-2C (FY-2C) in three infrared channels (IR1, 10.3-11.3μm; IR2, 11.5-12.5μm and WV 6.3-7.6μm) with 2864 samples collected by meteorologists in June, July, and August in 2007. The result shows that this method can distinguish and segment cloud patches of different types, and improves the traditional SRG algorithm by reducing the uncertainty of seeds extraction and regional growth.

Keywords : FY-2C, multi-channel satellite image, Artificial Neural Network, Seeded Region Growing, cloud patch, segmentation, cloud type.

I. Introduction

Clouds play an important role in the earth–atmosphere system. They significantly affect the heat budget by reflecting short-wave radiation (Hobbs 1981), and absorbing and emitting long-wave radiation (Hunt 1982). Cloud segmentation based on satellite images is vital for the parameter extraction of clouds patch and cloud track which are useful for numerous climatic, hydrologic and atmospheric applications. Therefore, accurate and automatic cloud segmentation has been a great interest of many scientists.

A wide variety of methods and algorithms are available to deal with the automatic segmentation of images (Fu and Mui 1981, Pal and Pal 1993). These methods can be broadly classified into four categories Zhu and Yuille 1996) : Edge - based techniques, (Region-based techniques, Deformable models, Global optimization approaches. Considering various shapes and blurry edges of cloud and the influence of underlying land surface, region-based techniques (Simpson et al. 1998, Shin et al. 1996) is suitable to detect clouds. For example, Seeded Region Growing algorithm (SRG) (Adams and Bischof 1994), exploit the spatial information of pixels and the formed regions will be homogeneous and connected. However, Conventional region-growing methods depend on the consistency of the image, and are sensitive to seed initialization and thresholds (Zucker 1976, Weihong et al. 2008). It is usually limited to regional model because cloud top bright temperature (TBB) varies greatly from different region and season.

with the development of computer, artificial intelligence has been applied to cloud segmentation (Okada et al. 2003, Schlüter and Heygster 2002, Peak and Tag 1994 Han et al. 2006). Most of those methods are based on texture parameters, which is hard to differentiate some clouds, such as thin cirrus and low-level clouds for they are often influenced by underlying land surface because of the closeness of their brightness temperature (TBB).Similarly, there are some new methods has been used, such as Markovian (Kussul et al.2005, Collet et al.2003), Hierarchies (Tilton 2006), mathematical morphology (Wang et al.2001), Bayes factors (Murtagh et al.2003), and so on(Din-Chang et al.2008, Papin et al.2002, Manizade et al.2006, Lim and Sagar 2008).

Almost all those method mentioned before usually use only one infrared image. Plenty information provided by other channels are wasted. Most important of all, there is no method to identify cloud patches of different types directly. To deal with this problem, this study tries to use Seeded Region Growing (SRG) as segmentation algorithm and pre-process the satellite images by using pixel-based classifiers based on multi-channel data. Because cloud classifiers can reduce the number of satellite image value from hundreds floats to several integers. In addition, it is possible to determine cloud types according to research objective.

As for pixel-based classifiers, a lot methods have been developed for remote-sensing instruments using various machine learning techniques, such as neural network (Key et al.1989), Bayesian methods

Author α : Tianjin Hydraulic Research Institute, Tianjin 300061, China. E-mail : liuyu950@126.com
Author σ : The Ocean Surveying & Mapping Institute of the Navy , Tianjin 300061, China. E-mail : chiduganggis@163.com.
Author ρ : Tianjin Hydraulic Research Institute, Tianjin 300061, China. E-mail : lxg@ tjhri.com
Author ω: National Satellite Meteorological Center, Beijing 100081,China. E-mail : shicx@cma.gov.cn.
Author ¥ : Tianjin Hydraulic Research Institute, Tianjin 300061, China. E-mail : zk@tjhri.com.

(Uddstrom et al.1999), clustering analysis or maximum likelihood (Li et al.2003), and fuzzylogic(Baum et al.1997). According to previous studies(Liu et al.2009), neural network of cloud classification works better in FY2C. In this study we examined the feasibility of using ANN to classify clouds from infrared data based on numerous samples collected by hands.

FY-2C is launched successfully in Beijing on October 19, 2004. It is the first Chinese-based operational geostationary meteorological satellite. One of the motivations for this work is to find out suitable techniques for the segmentation of clouds image in preparation for the upcoming launches of the FY-4 series.

Section Ⅱ introduces the FY-2C images and data, and provides a brief description of the segmentation methods. In Section Ⅲ the segmentation results of different cloud types are presented, and the capability of ANN on cloud classification are demonstrated in three aspects: time cost, robustness, precision. The discussions and summary are given in section Ⅳ.

II. Data and Methodology

a) Data

1) Satellite Data

FY-2C is positioned over the equator 105° E, and carries VISSR (Visible and Infrared Spin Scan Radiometer). Its nadir spatial resolution is 1.25 km for visible channel, and 5 km for infrared channels. Considering the remote sensing characteristics of cloud, automatic cloud classification system by which clouds in daytime and night can be compared, three infrared channels haven been chosen: IR1, 10.3-11.3μm; IR2, 11.5-12.5μm; WV 6.3-7.6μm.

2) Samples

In this study, 2864 samples are colleted by hands with the help of several experienced meteorologists. Special human-computer interactive software which is developed by Ph.D C.J.Yang in NSMC (National Satellite Meteorological Center in Beijing) in the Windows environment has been used. Those samples are major composed of that of June, July and August in 2007. They cover almost all the cases in the summer of 2007 and they can meet the needs of cloud classification model. The more detail description of cloud samples can be seen in Yu Liu(2009).

b) ANN-SRG Model Architecture

The ANN-SRG technique we proposed is an iterative process by which regions are merged starting from individual pixels, or another initial segmentation, and growing iteratively until every pixel is processed. It consists of two stages: the pre-processing by using an artificial neural network (ANN) to classify cloud, and cloud segmentation by using a seeded region growing algorithm (SRG). The scheme of the proposed algorithm is shown in Fig.1 and Fig.2. It can be described by the following steps:

1) ANN Cloud Classification Model

In order to simplify image analysis and computer vision for the subsequent segmentation step, this study pre-processed satellite image by classifying cloud of FY-2C data with ANN in three infrared channels (IR1, IR2,WV). Seven categories have been divided: sea, land, low-level clouds, midlevel clouds, thin cirrus, thick cirrus,multi-layer clouds, cumulonimbus. Its detail description of the ANN cloud classifier of FY2C can be seen in Yu Liu(2009).The schematic diagram of the result can be shown as Fig.2 (b).

2) Seeded Region Growing

The seeded region growing technique is an iterative process by which regions are merged starting from individual pixels, or another initial segmentation, and growing iteratively until every pixel is processed. It can be described by the following steps:

1) *Pre-processing* : Segment the entire image into pattern cells. Details are described in the previous part: *B(1). ANN Cloud Classification Model*.
2) *Seed extraction :* Scan image and find out a cell that hasn't been labeled. This step try to find out seed (CT_{seed}) whose cloud type is desired (CT_{seed}= $CT_{desired}$) and hasn't been labeled. As it is shown in Fig.2 (c), this method is not sensitive to seed positioning given a certain cloud type;
3) *Region growing :* Each pattern cell is compared with its neighboring cells to determine if they are the same type ($CT_{i,\ j}$= CT_{seed}). If they are same, merge the cells to form a fragment and update the property used in the comparison. Then Continue growing the fragment by examine all of its neighbors until no joinable regions remain. Label the fragment as a completed region. The schematic diagram of the result of ANN-SRG (Fig.2 (c)) shows that this method not sensitive to region growing principal.
4) *Unlabelled Parts Selection :* selects unlabelled parts to reapply steps 2), 3).

The four steps are iterated until the whole image is labeled.

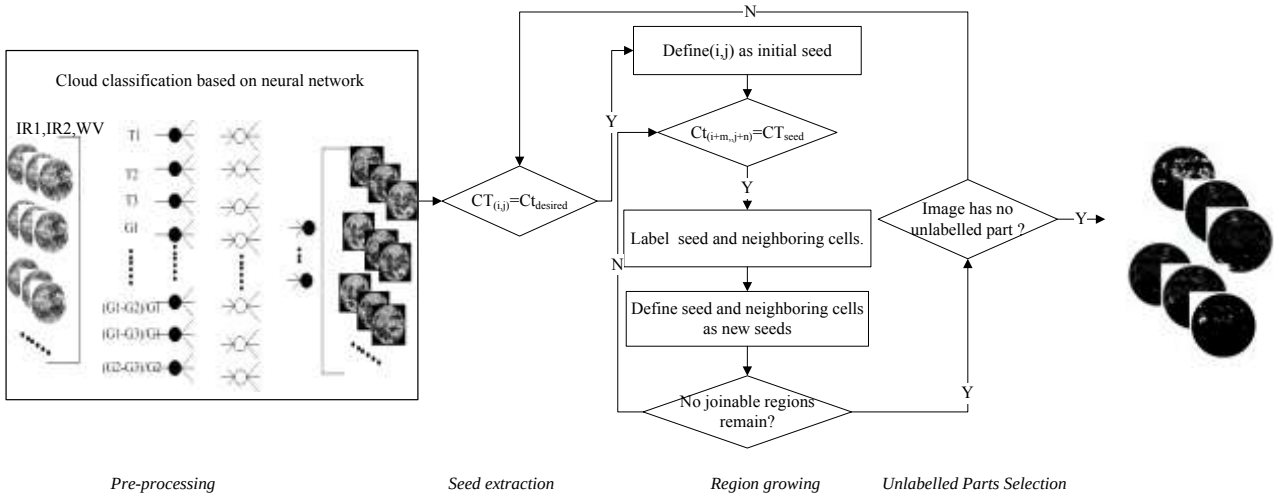
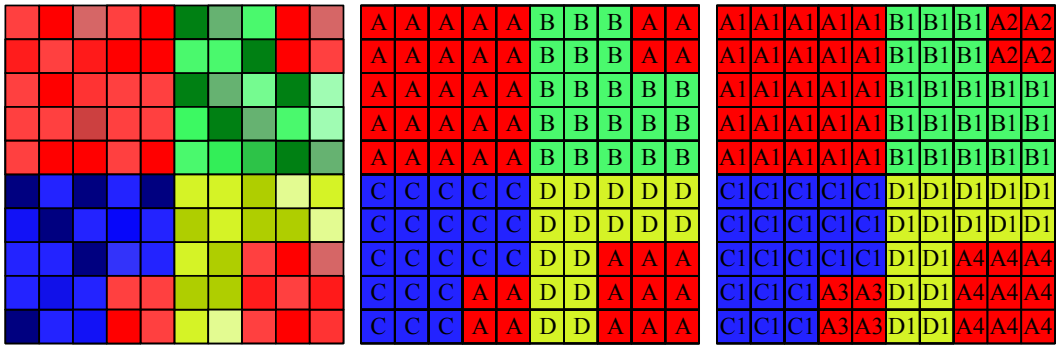


Fig.1: Configuration for the cloud segmentation; CT_{seed}-Cloud type of seed; $CT_{desired}$-Desired cloud type; $CT_{i,j}$-Cloud type of pixel (i, j).



(a) (b) (c)

Fig.2: Schematic diagram of the result of ANN cloud classification and ANN-SRG segmentation

(a) Original satellite map. Different color shows different cloud characteristics. (b) Result of ANN cloud classification. A, B, C and D are cloud types. (c) Result of ANN-SRG segmentation. Ai (A1 , A2…An) is the cloud patch i whose cloud types is A. B1, C1 and D1 are the first cloud patch whose cloud types are B, C and D respectively.

c) Evaluating Model Performance

The performance of the ANN-SRG can be judged by the three aspects: time cost, robustness, precision. The first one can be demonstrated by the training time consumed.

To analyze the robustness of the model, this study used different percentage (10%, 20%, 30%, 40%, and 50%) of validation cases, and performs the model under the existence of external disturbances (5%, 10%, 15% of fraud samples in training and validation phase). Because the uncertainty of the ANN-SRG method only comes from ANN classification as mentioned before, this study mainly evaluates the precision of ANN classifier. Its performance can be indicated by confusion matrix for the test data of cloud classifier, and some indexes, such as mean square error (MSE), normalized mean square error (NMSE), error (%), correlation coefficient. These indexes are defined as the following:

1) *Mean square error (MSE)*

$$MSE = \frac{\sum_{j=0}^{p}\sum_{i=0}^{N}(d_{ij} - y_{ij})^2}{N \times P} \quad (1)$$

Where d_{ij} is desired output for exemplar i at processing elements j; y_{ij} is network of output for exemplar i at processing elements j; N is number of exemplars in the data set; P is number of output processing elements.

2) *Normalized mean square error (NMSE):*

$$NMSE = \frac{P \times N \times MSE}{\sum_{j=0}^{p}\frac{N\sum_{i=0}^{N}d_{ij}^2 - (\sum_{i=0}^{N}d_{ij})^2}{N}} \quad (2)$$

Where dy_{ij} is renormalized network of output for exemplar i at processing elements j; dd_{ij} is renormalized desired output for exemplar i at processing elements j.

3) *Percent Error (%):*

$$Error = \frac{100 \times \sum_{j=0}^{p} \sum_{i=0}^{N} \frac{\left| dy_{ij} - dd_{ij} \right|}{dd_{ij}}}{N \times P} \quad (3)$$

4) *Correlation coefficient (Corr):*

$$r = \frac{\sum_{j=0}^{p} \sum_{i=0}^{N} (d_{ij} - \bar{d})(y_{ij} - \bar{y})}{\sqrt{\sum_{j=0}^{p} \sum_{i=0}^{N} (d_{ij} - \bar{d})^2 \sum_{j=0}^{p} \sum_{i=0}^{N} (y_{ij} - \bar{y})^2}} \quad (4)$$

Where $\bar{d}$ is desired output for exemplar; $\bar{y}$ is network of output for exemplar.

III. Experimental Results

The segmentation result of 6 types of cloud patches (Low cloud, Medium cloud, Thin cirrus, Thick cirrus, Multi-layer cloud, Cumulonimbus) are presented in *Fig.3*.

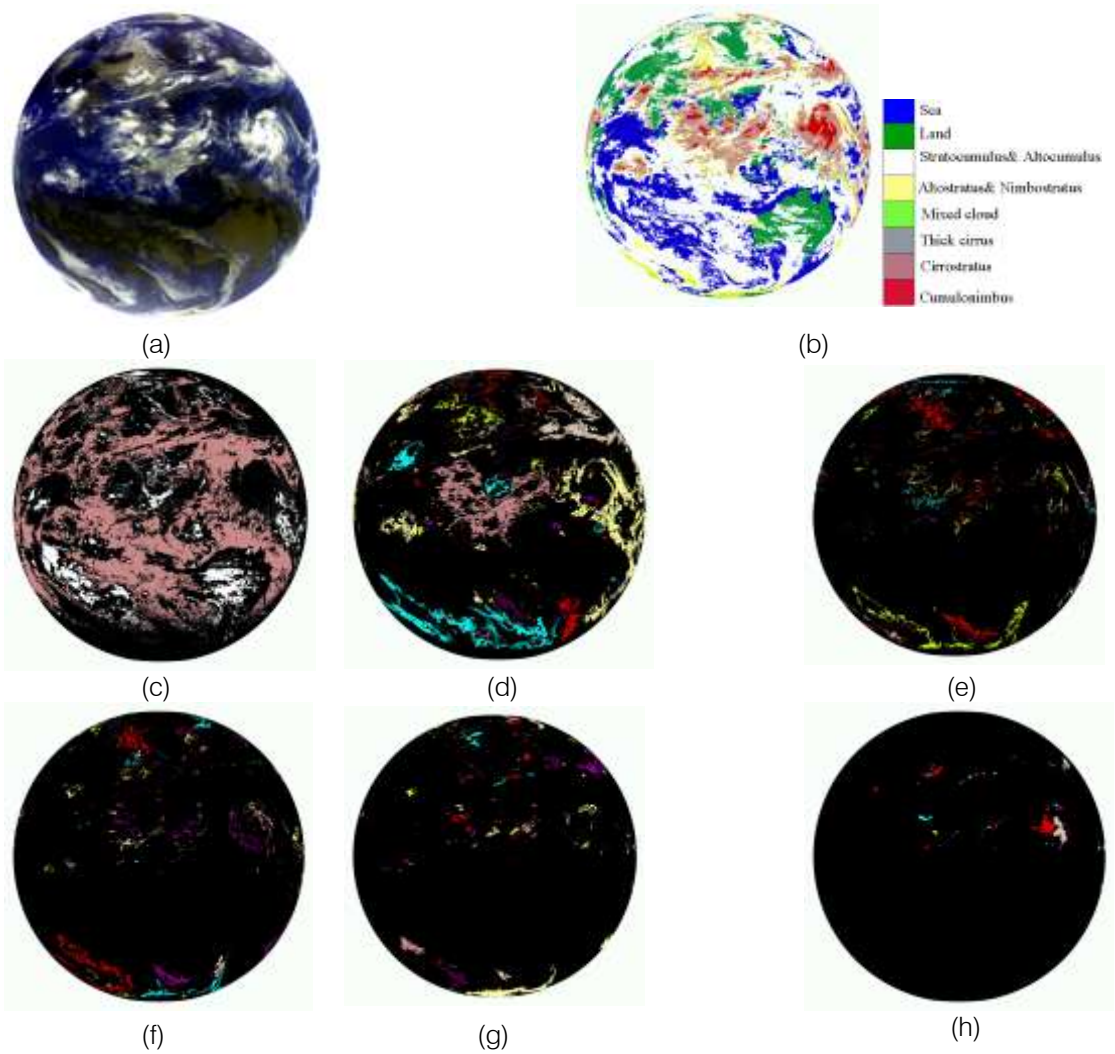


Fig.3: **(a)** pseudo-color composite map of Tbb of IR1, IR2 and WV(16: 00 UTC on 9 July2007); **(b)** cloud classification results of ANN ;**(c)** Segmentation result of low cloud; **(d)** Segmentation result of medium cloud; **(e)** Segmentation result of thin cirrus; **(f)** Segmentation result of thick cirrus; **(g)** Segmentation result of multi-layer cloud; **(h)** Segmentation result of cumulonimbus.

a) ANN-SRG Segmentation Precision

The segmentation precision is analyzed from the pixel level and the image level which is given with the comparison of traditional SRG.

1) Segmentation precision of pixel

The model errors is 8.87% in test period and the MSE, NMSE and correlation rates (Corr) of test and cross-examination of of ANN are 0.01, 0.02 0.99

respectively. It shows that ANN-SRG methods can classify and segment clouds well.

The confusion matrix in Table 1 demonstrates that the model can differentiate 87%- 98% of Low-level clouds, Midlevel clouds, Thin cirrus, Thick cirrus, Multi-layer clouds, Cumulonimbus accurately. The misjudgment always related to the similarity of clouds temperature, for example, it misjudged multi-layer clouds as cumulonimbus and thick cirrus which all of them has low temperature.

Table 1. Confusion matrix for the test data of cloud classifier

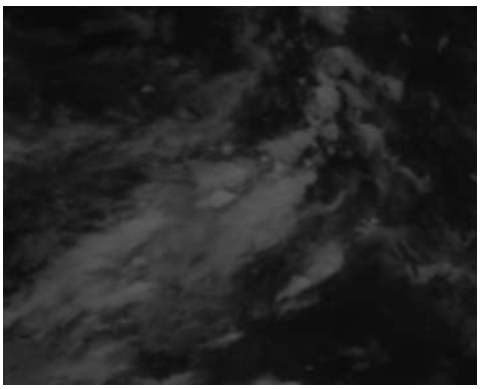
Classes	Sea	Land	Low-level clouds	Midlevel clouds	Thin cirrus	Thick cirrus	Multi-layer clouds	Cumulonimbus
Sea	1	0	0	0	0	0	0	0
Land	0	0.97	0	0.01	0.01	0.01	0	0
Low-level clouds	0.04	0.05	0.87	0.02	0.02	0	0	0
Midlevel clouds	0.01	0	0.02	0.92	0.05	0	0	0
Thin cirrus	0.01	0.01	0.02	0.01	0.93	0.02	0	0
Thick cirrus	0	0	0	0.01	0.02	0.92	0.05	0
Multi-layer clouds	0	0	0	0.03	0	0.01	0.90	0.06
Cumulonimbus	0	0	0	0	0	0.01	0.01	0.98

2) *Segmentation Precision of Image*

Fig.3 show that the configuration of cloud patches has smooth boundary, and low noise. They are close to the reality according to Meteorologists. The ANN-SRG method can detect different types of cloud well.

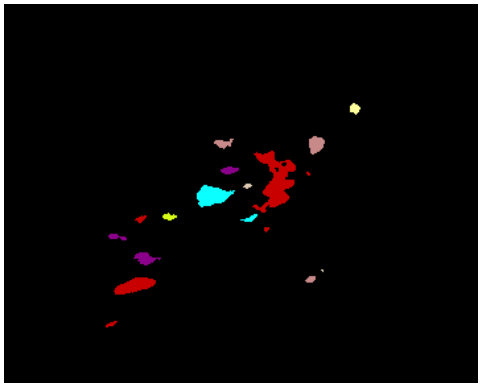
To compare the segmentation result of ANN-SRG with traditional SRG, this study chooses Mesoscale convective complexes (MCCs) case. Reasons for limit to the MCCs cases lie in the following aspects: firstly, MCCs are responsible for most of the warm-season rainfall (Augustine and Howard 1991); Secondly, According to the Physical mechanism of MCCs, it is composed of Cumulonimbus and Multi-layer clouds. The segmentation result by SRG and ANN-SRG can be compared; Thirdly, it is not possible to identify clouds with only one satellite channel data except MCCs whose TBB is less than -52°C according to its definition (Augustine and Howard 1991, Jirak *et al.*2003). Therefore, this study use -52°C as the growth principle of traditional SRG.

Segmentation result of a Mesoscale convective complexes (MCCs) case of the traditional SRG and ANN-SRG can be seen in Fig.4. It shows that ANN-SRG method can provide more detailed information by differentiating Cumulonimbus well from Multi-layer clouds. On the other hand, it improves the traditional SRG by reducing the uncertainty of seeds extraction and regional growth.



(a)

(b)



(c)

(d)

Fig.4: Segmentation result of the SRG and ANN-SRG(16: 00 UTC on 9 July2007).**(a)** Infrared map of Channel 1; **(b)** Segmentation result of MCCs by SRG; **(c)** Segmentation result of Cumulonimbus by ANN-SRG; **(d)** Segmentation result of Multi-layer cloud by ANN-SRG

b) Robustness of ANN-SRG

The robustness of the ANN-SRG model for different validation samples and some fraud samples can be seen in Table 3. It clearly suggests that error of the model is about 10% in test period with 5-50% (Table 2 Model[1])validation samples and 5-15% (Table 3 Model[1])fraud samples for validation and the training.

Table 2. The result of cross-examination and tests of cloud classification model for different validation samples and for some fraud samples

Percent	Cross-examination				Test			
	MSE	NMSE	Corr	Error(%)	MSE	NMSE	Corr	Error(%)
Model[1]								
5%	0.01	0.03	0.99	8.92	0.01	0.02	0.98	9.18
10%	0.01	0.02	0.99	9.35	0.01	0.02	0.99	8.87
20%	0.01	0.02	0.99	9.47	0.01	0.02	0.99	8.94
30%	0.01	0.02	0.99	9.78	0.01	0.02	0.99	8.99
50%	0.01	0.02	0.99	9.52	0.01	0.02	0.99	11.91
Model[2]								
5%-5%	0.01	0.02	0.10	0.99	0.01	0.03	0.99	9.52
10%-5%	0.01	0.03	0.10	0.98	0.01	0.03	0.99	9.62
15%-5%	0.01	0.02	0.10	0.98	0.01	0.03	0.98	8.47
5%-10%	0.01	0.03	0.10	0.98	0.01	0.03	0.98	10.40
10%-10%	0.01	0.04	0.10	0.98	0.01	0.03	0.98	10.77
15%-10%	0.01	0.04	0.10	0.98	0.01	0.03	0.98	10.96
5%-15%	0.01	0.03	0.10	0.98	0.01	0.03	0.98	12.10
10%-15%	0.01	0.03	0.10	0.98	0.01	0.03	0.98	12.00
15%-15%	0.01	0.04	0.10	0.98	0.01	0.04	0.98	12.58

The classification result of ANN for different validation samples. The sample for all models in this table consists of 2864cases. For the 5% model, the training sample is 95% cases (2721) and the validation sample is 5% cases (143). The others also correspond to this system so that the 15% model has 15 %(429) cases in its validation sample.

The classification result of ANN with some fraud samples. The model trains on the 2292 cases and uses the last 572 as a validation. For the 5%-5% case, the training sample is 5% fraud cases (114) and 95% non-fraud cases (2178). The validation sample is 5% fraud cases (28) and 95% non-fraud cases (544). The others also correspond to this system so that the 5%-15% model has 5% fraud cases (114) in training sample and 15 %(85) fraud cases in its validation sample.

c) Execution Time

We have measured the execution times of SRG and ANN- SRG (Table 3). It shows that the SRG consumes only 66 seconds to segment MCCs, while ANN-SRG consumes about 5 minutes for each cloud patches. However, the results of ANN-SRG are acceptable because FY-2C real time infrared images are collected hourly. Therefore, there is no obvious time limit for the proposed method.

The results also demonstrate that the pre-processing period of ANN- SRG occupies the most of time during segmentation process. Further improvement should be carried out. The execution time of different types of cloud is quite different. The fastest one is the segmentation of cumulonimbus, because of it its low frequency of occurrence on satellite image compared to other types of clouds.

Table3. Execution time FY-2C segmentation system(S)

Type			Execution time
ANN-SRG	ANN		225
	SRG	Low cloud	62
		Medium cloud	50
		Thin cirrus	53
		Thick cirrus	54
		Multi-layer cloud	52
		Cumulonimbus	43
	Total		539
SRG		MCCs	66

IV. Conclusion

The study has presented a new segmentation scheme (ANN-SRG) to detect cloud patches of different type directly by the combination of artificial intelligent and traditional segmentation methods, and applied it to FY-2C in three infrared channels (IR1, 10.3-11.3μm; IR2, 11.5-12.5μm and WV 6.3-7.6μm). The results show that the ANN-SRG method can take advantage of multi-channel satellite data and segment cloud patches of different types which can't be achieved by traditional segment method. It improves traditional SRG by reducing the uncertainty of seeds extraction and regional growth, because of using ANN cloud classifiers to reduce the number of infrared image value from hundreds floats to several integers.

The segment results of cloud patches obtained so far have been encouraging. Region boundaries have good correspondence with the contours of the cloud patches. However, in reality ANN-SSR method might not be able to achieve such good results as demonstrated in this study. There are mainly two reasons to constrain the further application of the current model. First, ANN-SSR method is sensitive to pre-processing result by ANN cloud classification which greatly depend on expert experience. Moreover, the infrared top brightness temperature (TBB) varies greatly in different region and seasons because of the influence of underlying land surface. To truly establish an effective and efficient cloud classification algorithm for the FY-2C satellite, more work needs to be carried out to analyze whether it is possible to build a cloud classification method independent of regions and seasons. Second, due to the relative complexity of ANN cloud classifier and series algorithm, the execution time of ANN-SSR is a little much, and its takes about 9 minutes for the cloud segmentation of each FY-2C image. Further improvement should be made, such as more quick cloud classifier and parallel algorithm.

Acknowledgements

Special thanks go to Pro.Xia,J.,and Ph.D Hong, Y.for their contributions to the manuscript.

References Références Referencias

1. Adams,R., and Bischof, L. (1994). "Seeded Region Growing. "*IEEE T. Pattern Anal.,* 16(6), 641-647.
2. Baum, B. A., Tovinkere, V., Titlow, J., and Welch, R. M.(1997)."Automated cloud classification of global AVHRR data using a fuzzy logic approach." *J. Appl. Meteor.*, 36, 1519-1540.
3. Collet, C., Louys, M., Obero, A., Bot, C., and Ieee, I. (2003)."Markov model for multispectral image analysis: Application to small magellanic cloud

segmentation."*IEEE International Conference on Image Processing, Ieee: Barcelona, Spain,* 2003, 953-956.
4. Tseng, D., Tseng, H., and Chien, C. (2008)."Automatic cloud removal from multi-temporal SPOT images."*Appl. Math.Comput.*,205, 584-600.
5. Fu,K., and Mui, J. (1981)."A survey on image segmentation."*Pattern Recogn.*, 13, 3-16.
6. Hobbs, P.V. (1981).*Clouds: Their formation, optical properties and effects*, Deepak, A., Eds., Academic Press: New York, 1981, P.285.
7. Hunt, G.E. (1982)."On the sensitivity of a general circulation model climatology to changes in cloud structure and radiative properties."*Tellus*, 34, 29-38.
8. Key, J., Maslanik, J.A., and Schweiger, A.J. (1989)."Classification of merged AVHRR/3 and SMMR Arctic data with neural networks."*Photogramm. Eng. Rem. Sens.,* 55, 1331-1338.
9. Kussul, N., Korbakov, M., Kravchenko, A., and Shelestoy, A. (2005)."Cloud mask extracting from Meteosat data with use of parallel Markovian segmentation algorithm."*2005 IEEE Intelligent Data Acquisition and Advanced Computing Systems:* Technology *and Applications*, 204 - 207.
10. Li, J., Menzel, W. P., Yang, Z., Frey, R. A., and Ackerman, S. A. (2003). "Highspatial- resolution surface and cloud-type classification from MODIS multispectral band measurements."*J. Appl. Meteor.,* 42(2), 204-226.
11. Lim, S. L., and Sagar, B. S. D. (2008)."Cloud field segmentation via multiscale convexity analysis."*J. Geophys.Res.*, 113(D13208), 17.
12. Liu, Y., Xia, J., Shi, C.X., and Hong, Y. (2009)."An Improved Cloud Classification Algorithm for China's FY-2C Multi-Channel Images Using Artificial Neural Network."*Sensors*, 9(7), 5558-5579
13. Manizade, K. E., Spinhirne, J. D., and Lancaster, R. S. (2006)."Stereo cloud heights from multispectral IR imagery via region-of-interest segmentation." *IEEE T. Geosci.Remote*, 44(9), 2481-2491.
14. Murtagh, F., Barreto, D., and Marcello, J. (2003)."Decision boundaries using Bayes factors: The case of cloud masks."*IEEE T. Geosci.Remote*, 41(12), 2952-2958.
15. Okada, Y., Waizumi, Y., Kato, N., and Nemoto, Y. (2003)."A study of automatic image segmentation using neural networks."*Record of Electrical and Communication Engineering Conversazione Tohoku University*, 72, 352-353.
16. Pal, N., and Pal, S. (1993)."A review on image segmentation techniques."*Pattern Recogn.,* 26,1277–1299.
17. Papin, C., Bouthemy, P., and Rochard, G. (2002)."Unsupervised segmentation of low clouds from infrared METEOSAT images based on a contextual spatio-temporal labeling approach."*IEEE T. Geosci.Remote*, 40(1), 104-114.
18. Peak, J. E., and Tag, P. M. (1994)."Segmentation of satellite imagery using hierarchical shareholding and neural networks."*J. Appl. Meteor.*, 33(5), 605-616.
19. Schlüter, N., Heygster, G. (2002)."Remote sensing of Antarctic clouds with infrared and passive microwave sensors."*Meteor. Z.*, 11(1), 21-36.
20. Tilton, J. C. (2006)."Cloud Mask Generation for MODIS Utilizing Hierarchical Segmentation."*IEEE International Geoscience and Remote Sensing Symposium (IGARSS),2006*, 3708-3711.
21. Uddstrom, M. J., Gray, W. R., Murphy, R., Oien, N. A., and Murray, T. (1999)."A Bayesian cloud mask for sea surface temperature retrieval."*J. Atmos. Ocean Tech.,* 16, 117-132.
22. Wang, P., Xue, J. T., Liu, Z. G., Liu, H. Z., and Tang, G. S. (2001)."Self-adaptive segmentation for infrared saterllite cloud image."*Conference on Image Extraction, Segmentation, and Recognition,* 2001, 388-393.
23. Weihong, C., Zequn, G., and Zhiyi, Z. (2008)."An improved region growing algorithm for image segmentation."*In 2008 International Conference on Computer Science and Software Engineering, Wuhan, Hubei, China*, 93-96
24. Yhann, S. R., and Simpson, J. J. (1995)."Application of neural networks to AVHRR cloud segmentation."*IEEE T. Geosci.Remote*, 33(3), 590-604.
25. Zhu, S. C., and Yuille, A. (1996)."Region competition: Unifying snakes, region growing, and Bayes / MDL for multiband image segmentation."*IEEE Trans.on Pattern Anal.Mach.Intell.*, 18(9), 884–900.
26. Zucker, S. W. (1976)."Region Growing: Childhood and Adolescence."*C Comput. Vis., Graph. Image Process,* 5(2),382-399.

Global Journal of Computer Science and Technology
Volume 12 Issue 7 Version 1.0 April 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

A Temporal Ontology for Reasoning about Actions

By Fatiha Mamache
Universté des Sciences et de la technologie Houari Boumediéne. Alger, Algérie

Abstract - In this paper, our work is devoted to systematic study of actions theories by using a logical formalism based on a first order language increased by operators whose main is to facilitate the representation of causal and temporal relationships between actions and their effects as well as causal and temporal relationships between actions and events. In Allen and Mc-Dermott' formalisms, we notice that notions of past, present and future do not appear in the predicate Ecause. How to affirm that effects don't precede causes? To use the concept of temporality without limiting themselves to intervals, we enrich our language by an operator defined on time-elements Our formalism avoids an ambiguity like: effect precedes cause. The originality of this work lies in proposal for a formalism based on equivalence classes. We also defined an operator who allows us to represent the evolutions of the universe for various futures and pasts. These operators allow to represent the types of reasoning which are prediction, explanation and planning. we propose a new ontology for causal and temporal representation of actions/events. The ontology used in our formalism consists of facts, events, process, causality, action and planning.

Keywords : *Artificial Intelligence, Description Logic, Knowledge Representation, Reasoning on the Actions, Spatio-Temporal Logic, Temporal Logic.*

GJCST Classification: *K.4.4*

Strictly as per the compliance and regulations of:

A Temporal Ontology for Reasoning about Actions

Fatiha Mamache

Abstract - In this paper, our work is devoted to systematic study of actions theories by using a logical formalism based on a first order language increased by operators whose main is to facilitate the representation of causal and temporal relationships between actions and their effects as well as causal and temporal relationships between actions and events. In Allen and Mc-Dermott' formalisms, we notice that notions of past, present and future do not appear in the predicate Ecause. How to affirm that effects don't precede causes? To use the concept of temporality without limiting themselves to intervals, we enrich our language by an operator defined on time-elements Our formalism avoids an ambiguity like: effect precedes cause. The originality of this work lies in proposal for a formalism based on equivalence classes. We also defined an operator who allows us to represent the evolutions of the universe for various futures and pasts. These operators allow to represent the types of reasoning which are prediction, explanation and planning. we propose a new ontology for causal and temporal representation of actions/events. The ontology used in our formalism consists of facts, events, process, causality, action and planning.

Keywords : Artificial Intelligence, Description Logic, Knowledge Representation, Reasoning on the Actions, Spatio-Temporal Logic, Temporal Logic.

I. Introduction

The temporal reasoning consists to formalize the notion of time and to provide means to represent and reason on the temporal aspects of knowledge. To describe the properties of the good performance of applications, temporal logics are formalisms well adapted , in particular by their capacity to express the scheduling of actions/events in time.

Classic logics are unsuited to temporal reasoning. One of the weaknesses of these logics is that the material implication takes account neither of temporal scheduling between causes and effect, nor of monotony of causal reasoning.

The causal reasoning is a non monotonous temporal reasoning. Concept of cause is usually used in daily life, we frequently attribute to people and to objects a causal capacity compared to the events. The human use their knowledge on relations on causes/effect type to reason on current situations of the life and to make decisions which generally determine the choice of actions to carry out to reach desirable effects or to avoid undesirable effects.

Author : LRIA Laboratoire de Recherche en Intelligence Artificielle. Faculté d'Electronique et d'Informatique. Département d'Informatique, USTHB, BP 32 El Alia, Bab Ezzouar, Alger.
E-mail : amamache@usthb.dz

Temporal logics having retained researchers attention are Allen and Mc-Dermott's logics . They are the most important formalisms of temporal representation. The time representations can be characterized by the primitive objects which they consider. Allen developed a temporal motor specialized (time specialist) to manage relations between temporal aspects of knowledge and on this basis he conceived a temporal logic. The Allen temporal motor's role is the management of relations between the intervals. Its ontology is constitute of properties, events and process.

Mc-Dermott proposed a formalism of causality, action and planning. For causality, he mentioned the qualification problem of a cause and the persistence problem of a fact. He pointed out that a solution of these problems is in a good formalization of the non monotonous reasoning.

In Allen and Mc-Dermott formalism's, we notice that notions of past, present and future do not appear in the predicate Ecause. How can one know if Ecause $(p, e_1, e_2, r, d_1, d_2)$ means that the event e_1 is always followed event e_2 after a time included in the interval (d_1, d_2), occurred in the past, present or future? How to affirm that effects preceding step causes?

We are interested by an agentive design of causality, closely related to the concept of action whose modelling must include two fundamental aspects: - Temporal aspect at the representative level (the cause must precede the effect) and, - Non monotonous aspect on the functional level of the causal relations (an effect must have a cause). The design of adopted causality is from the formalization of the causal and temporal reasoning.

Our work is devoted to systematic study of actions theories by using a logical formalism based on a first order language increased by operators whose main aim is to facilitate the representation of causal and temporal relationships between actions and their effects as well as causal and temporal relationships between actions and events. Our formalism avoids an ambiguity like: effect precedes cause.

The originality of this work lies in proposal for a formalism based on equivalence classes. We also defined an operator who allows us to represent the universe evolutions for futures and passed varied. These operators allow to represent the types of reasoning which are prediction, the explanation and planning.

We propose a new ontology for causal and temporal representation of actions/events. The ontology used in our formalism consists of facts, events, process, causality, action and planning.

The paper is organized as follows: In the next section, we establish the formal background that will be used throughout this paper. In section 3, we propose a new ontology for causal and temporal representation of actions/events. The ontology used in our formalism consists of facts, events, process, causality, action and planning. In section 4, we define syntax and semantics of our temporal logic $\mathcal{L}_C$. We also define the valuation in the following cases:

– Case of the effects/events which require the realization of several actions at the same time. In this case, we represent the set of the actions occurred at the same time by the equivalence class of an action which is the representative of the class.
– Case of an action which is repeated in different time-element (process). We represent the set of the time-elements by the equivalence class of a time-element which is the representative of the class.
– Case of the competitive actions. We have two possibilities for the choice.

(i) Temporal choice
(ii) Economic choice

Section 5 is devoted to completude and in section 6 we conclude with a general idea of researches on actions theory.

II. Language, Notation and Terminology

a) Introduction

Within the framework of the formalization of an approach symbolic system for the temporal and causal reasoning, and inspired by work of [Allen, 84,][1], [McDermott, 82][3] and [Kayser and Mokhtari, 98][4], we propose a temporal causal formalism to reason on events and actions [Mamache, 2010][5].

The language is composed of two nival:

– To represent static information, the first level consists of a first order language with equality.
– the second level includes the predicates with temporal variables to represent dynamic information.

- Connectors:$\neg, \vee, \wedge, \supset$ and $\supset_c$ (causal implication)
- Two signs of quantification noted $\exists$ (existential quantifier) and $\forall$ (universal quantifier).
- A symbol of equality, which we will note $\equiv$ to distinguish it from the sign $=$.
- A countable infinite collection of propositional variable.
- A set of operational signs or symbols functional.
- Three unary temporal operators: P_k (past), F_k (future), and P_0 (present).
- The expressions are the symbol strings on this alphabet.
- The set of the formulas noted Φ is by definition the smallest set of expressions which checks the following conditions :
  - Φ contains the propositional variables.
  - A set of elements called symbols of individuals.
  - If A and B are elements of Φ it is the same for $\neg A$ and $A \supset_c B$.
  - If A is an element of Φ it is the same for $P_k A$, $F_k A$ and $P_0 A$.

To introduce causality J. Allen [1][2] uses the following formula:

Ecause (p_1,i_1,p_2,i_2). It expresses, thus, the fact that p_1 which occurs in i_1 caused the event p_2 which occurs in i_2.

Like Allen, we use the predicate Ecause to express that an action a is the cause of an event e.

b) Causality Atemporal Representation

In the following e designate an effect of the action a or an event caused by the action a.

To express that an action a is the cause of an event e or an effect of a, as Allen, we use the predicate Ecause(a;e).

– If a is not the cause of e, we use $\neg$ Ecause $(a; e)$. In this case, the realization of e is due to another action.
– If a is the cause of the not realization of e, we use Ecause $(a; \neg e)$.
– If e is not realized because the action a is not executed, we use Ecause $(\neg a; \neg e)$. In this case a is a direct cause of e.

The actions seem first argument of the Ecause predicate.

The case where several actions $a_1,a_2,...,a_m$ are the cause of the same effect or a single event is expressed by the formula:

Ecause $(a_1,a_2,...,a_m; e)$ defined by

Ecause $(a_1,a_2,...,a_m; e) \equiv$ Ecause$(a_1;e) \wedge ... \wedge$ Ecause $(a_m; e)$

where $a_1,a_2,...,a_m$ are the atemporal expressions of actions type.

Definition 2.1 : Actions $a_1,a_2,...,a_m$ are said to be direct cause of an event e if as soon as one of these actions is not carried out, the event is not executed.

This formula can be expressed as :

Ecause $(a_1,a_2,...,a_m;e) \equiv ((\exists k)(\neg a_k \supset_c \neg e))$, if $a_1,a_2,...,a_m$ are direct causes of e.

Example 2.2 *:* Ecause(prepare one's paper, travelling, ..., communicate) ≡ (¬ travelling) $\supset_c$ (¬ communicate) $\vee$ (¬ no prepare paper) $\supset_c$ (¬ communicate)$\vee$...

c) Causality Temporal Representation

If a is a temporal expression of action type we use the following formulas :

- $t \cdot a$ if a is produced in the past at the element of time t.
- $a \cdot t$ if a it happens in the future at the element of time t.

We will keep the same notations in the case of an event (or effect) e :

- $e \cdot t$ for the future.
- $t \cdot e$ for the past.

Example 2.3

a) *Colloquium · May, means:* the colloquium will be held in May.
b) *May · Colloquium, means:* the colloquium was held in May.

If a is an action carried out in t' then the predicate Ecause $(a.t'; e.t)$ expresses the fact that a carried out in t' is the cause of e true in t.

This notation avoids an ambiguity like:an action which will occur in the future in t' is the cause of the event e which occurred in the past in t (the effect precedes the cause). Thus the expression Ecause $(a.ta'; t.e)$ does not have a 'sense'.

An action can be instantaneous as it can be carried out during in a certain interval of time [Knight, 98][6],[Knight, 97] [7]. Consequently, the points and the intervals are necessary to express the execution time of an action.

We call time-element an interval or a point of time. Therefore, an action operates during a time-element [Knight, 98][6],[Knight, 97] [7].

Definition 2.4 : A point of time T is an instantaneous state of the universe defined by a subset of true proposals in a certain date and by this date.

This subset is the result of a causal relation. The set of points of time is noted P.

Definition 2.5 : A time-element is an interval or a point of time. An action thus operates during a time-element t. If a is an instantaneous action then t is a point of time. If a durative then t is an interval.

Definition 2.6 : Let T a nonempty set of time-elements, T is the union of two sets P and I, I is a set whose elements are intervals and P a set whose elements are points of time [Birstougeff, Ligozat, 89][9].

Definition 2.7 : Let T a nonempty set of time-elements and A a set of actions. $A.T$ is defined as being the set of elements $a.t$ where a is a temporal expression of action type which will be carried out in the future in the time-element t.

Definition 2.8 : Let T a non empty set of time- elements, A a set of actions, $A{\cdot}T$ the set of elements $a{\cdot}t$ and Dur_F; an application from $A{\cdot}T$ to IR_+ defined by [Knight, 98][6],[Knight, 97] [7].:

$$\begin{cases} Dur_F\,(a \cdot t) = 0 & \text{if } a \text{ is an instantaneous action, thus, } t \text{ is a point of time .} \\ Dur_F\,(a \cdot t) > 0 & \text{if } a \text{ is a durative action, thus, } t \text{ is an interval.} \end{cases}$$

Definition 2.9 : Let T a non empty set of time- elements, A a set of actions, $T{\cdot}A$ the set of elements $t{\cdot}a$ and Dur_P; an application from $T{\cdot}A$ to IR_+) defined by [Knight, 98][6],[Knight, 97] [7].:

$$\begin{cases} Dur_P\,(t \cdot a) = 0 & \text{if } a \text{ is an instantaneous action, thus, } t \text{ is a point of time .} \\ Dur_P\,(t \cdot a) > 0 & \text{if } a \text{ is a durative action, thus, } t \text{ is an interval.} \end{cases}$$

The primitive temporal entities are time-elements.

Case where several actions $a_1,a_2,...,a_m$ are the cause of the same effect or a single event.
If $a_1,a_2,...,a_m$ are the temporal expressions of actions type carried out respectively in $t_1,t_2,..,t_m$, we use the formula :

Ecause $(a_1.t_1,a_2.t_2,...,a_m.t_m; e.t)$ ≡ Ecause $(a_1.t_1; e.t) \wedge ... \wedge$ Ecause $(am.t_m; e.t)$.

Example 2.10 : Ecause (January. prepare one's paper, send paper. April, ..., travelling. 15May; Communicate.18 June) ≡ Ecause (January. prepare one's paper; communicate.18 June) ∧...∧ Ecause (travelling.15 May;communicate.18 June).

Example 2.11 : The fact of travelling on Monday to communicate on Wednesday can be expressed as follows :

a) Ecause(travelling. Monday; communicate .Wednesday) expresses: the agent will travel on Monday in order to communicate on Wednesday.
b) Ecause(Monday. traveling; communicate. Wednesday) expresses: the agent travelled on Monday in order to communicate on Wednesday.
c) Ecause(Monday. travelling; Wednesday. communicate) expresses: the agent travelled on Monday and communicated on Wednesday.

The action of travelling has the effect communication on Wednesday, as effects do not precede action so we cannot have:

Ecause (travelling. Monday; Thursday .communicate)

An action a can be primitive as it can be complex. In the case of a complex action, to express that the actions $a_{i1},...,a_{is}$ carried out in $t_{i1},...,t_{is}$ (precondition) are the cause of a_i realized in t_i and this one will cause the effect (or event) e carried out in t we define:

Definition 2.12

$$\text{Ecause } (a_i.t_i;t.e). \equiv \text{Ecause } (a_{i\ 1}.t_{i1},a_{i2}.t_{i2},\ldots,a_{is}.t_{is};e.t)$$
$$\equiv \text{Ecause } (a_{i1}.t_{i1}) \wedge \text{Ecause } (a_{i2}.t_{i2}) \wedge \ldots \wedge \text{Ecause } (a_{is}.t_{is};e.t)$$
$$\equiv \bigwedge_{j=1}^{j=s} \text{Ecause } (a_{ij}.t_{ij};e.t).$$

The basic sets are:

a) A a set of the actions,
b) E a set of the events/effects, and
c) T a set of the time-elements.

To represent the connection which links a_n to its effect/events, we define the following application :

Definition 2.13

$$\zeta_{ev} : A \rightarrow E$$
$$a \mapsto \zeta_{ev}(a) \equiv e.$$

If event/effect requires several actions $a_1,a_2,\ldots,a_m$, we define:

Definition 2.14

$$\zeta_{ev} : A \times A \times \ldots \times A \rightarrow E$$
$$a_1,a_2,\ldots,a_m \mapsto (a_1 \wedge a_1 \wedge \ldots \wedge a_m) \equiv e.$$

The function which associates to an action a the time-element t_a in which it is carried out is defined as follows:

Definition 2.15

$$f_a : A \rightarrow T$$
$$a \mapsto f_a(a) \equiv t_a.$$

We defines the function which associates to an event e the time-element t_e of which it is carried out by:

Definition 2.16

$$f_e : E \rightarrow T$$
$$e \mapsto f_a(e) \equiv t_e.$$

An action causes an event/effects after a time allowed Δt. $t_e = t_a + \Delta t$. If $\Delta t = 0$ the action a and the event e occur at the same time.

Definition 2.17 *:* The set of the time-elements is projected on the axis of reals by a function date which associates to any element t of T its date noted d_t.

$$f_e : E \rightarrow T$$
$$e \mapsto f_a(e) \equiv t_e.$$

If $\Delta t = 0$ then $d_t \ \ a \ = d_{te}$.

Definition 2.18 *:* An Annal of time has [line of time for Kayser and Mokhtari, 98] [4] is a succession of time-elements Tt representing an evolution of the universe.

A point of time of the succession answers the rule 'there are no effects without cause', it is the result of a relation 'cause to effect'.

An annal of time is a convex unit, completely ordered in bijection with the axis of reals.

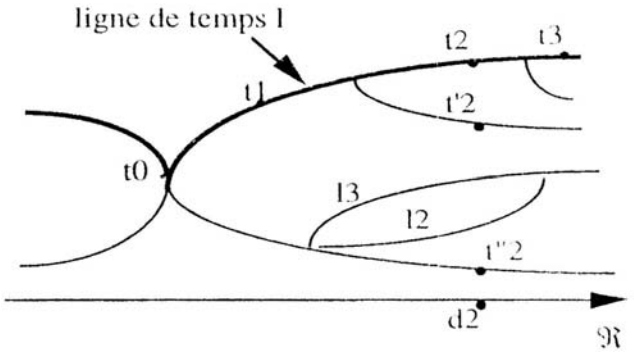


Figure 1: An Annal of time

III. A New Ontology to Represent Causal and Temporal Relationships Between Actions and Events/Effects

The ontology used in our language consists of effects, events and process.

a) Fact

A fact p is true in a point of time or interval. The notation True (p,t) expresses that the fact p is true in the time-element t.

b) Event

An event is carried out in a time-element. In the case of an interval, the events are true in the intervals where they are defined. They are not defined in the subintervals.

c) Processus

The processes are defined on intervals. If a process is true on an interval, it is true also on all subintervals of this interval.

d) Causality

An event causes another event.

If $e_1, e_2, \ldots, e_m$ are a temporal expressions of events type carried out respectively in $t_1,t_2,\ldots,t_m$, the formula:

a) Ecause $(e_1.t_1,e_2.t_2,\ldots,e_m.t_m;\ e.t) \equiv$ Ecause $(e_1.t_1;\ e.t) \wedge \ldots \wedge$ Ecause $(e_m.t_m;\ e.t)$; $e.t)$ expresses that the events $e_1,e_2,\ldots,e_m$ which will be realized respectively in $t_1,t_2,\ldots,t_m$ will cause the event e which will take place in the time-element t.

b) Ecause $(t_1.e_1,t_2.e_2,\ldots,t_m.e_m;\ e.t) \equiv$ Ecause $(t_1.e_1;\ e.t) \wedge \ldots \wedge$ Ecause $(t_m.e_m,e.t)$ expresses that the events $e_{11},e_2,\ldots,e_m$ which are realized respectively in $t_1,t_2,\ldots,t_m$

will cause the event e which will take place in the time-element t.

c) Ecause $(t_1.e_1,t_2.e_2,...,t_m.e_m;\ t.e)$ ≡ Ecause $(t_1.e_1;\ t.e)$ ∧...∧Ecause$(t_m.e_m,t.e)$ expresses that the events $e_{11},e_2,...,e_m$ which are realized respectively in $t_1,t_2,...,t_m$ will cause the event e which will take place in the time-element t.

e) Action and Planning

We are still inspired by Allen's work, an action is carried out by an agent and it produces an event/effect. Planning consists to defining a sequence of actions to be carried out by an agent to solve a general or specific problem. In addition to the construction of a sequence of obligatory or optional actions, J.Allen uses the concept of belief and intentionality. He proposes the following principles:

a) An agent S carries out an intentionally action a if and only if: - the agent carries out the action in a given interval;
b) The action belongs a plan that the agent had been committed carrying out during a given time interval.

J.Allen [1] is limited to the intervals.To use concept of temporality in planning and without limiting themselves with the intervals, we enrich our language by an operator noted ⊕. Our operator is defined on time-elements.

Definition 3.1: $t_1 \oplus t_2$ is defined if there are two actions a_1 and a_2 taking place in t_1 and t_2 respectively and which are the cause of an event (or effect) e carried out in a point of time t.

This operator has the following characteristics:

- ⋆ The operator ⊕ is internal if $t \in T$ (the agent must act so that the event or effect takes place in time-element t belonging to T).
- ⋆ The operator is commutative if the order of the actions does not intervene (the agent is free to start with any action). We denote: $t_1 \oplus t_2 \equiv t_2 \oplus t_1$.

J. A. Pinto [Pinto, 94] [8] established in his thesis a relation between events, actions and situations but he finds it more convenient to establish a relation between events, actions which occur for the realization of these events and the time when they are carried out. In our approach, we establish a relation between events, actions who occur for the realization of these events and time when they are carried out.

To express the fact that the actions $(a_1,a_2,...,a_m) \in A \times ... \times A$ which take place respectively, $t_1, t_2,..., t_m$ are the cause of an event e carried out in $t \in T$, we define the following diagram:

Definition 3.2 :

$$\begin{array}{ccc} A \times A \times ... \times A & \xrightarrow{\zeta_{ev}} & E \\ \downarrow \varphi & & \uparrow \phi \\ T \times T \times ... \times T & \longrightarrow & T \end{array}$$

where $\varphi(a_1,a_2,...,a_m) = (f_1(a_1), f_2(a_2),...,f_m(a_m))$, $f_i(a_i) = t_i \ \forall\ i \in 1,2,...,m$ and h a function defined as follows:

$$h : T \times T \times ... \times T \rightarrow T$$

$$h(t_1,t_2,\bullet,\bullet,\bullet,t_m) = t_1 \oplus t_2 \oplus \bullet\bullet\bullet \oplus t_m \equiv t.$$

h is defined if there exist actions $a_1,a_2,\cdot,\cdot,\cdot,a_m$ carried out respectively in $t_1,t_2,\cdot,\cdot,\cdot,t_m$ which gave place to e realized in t.

The intervening order of the actions in some events plays a significant role; like carrying out an action before another, reproduction of an action (process) or to carry out several actions at the same time. This led us to introduce operators on the actions. These operators define constraints over time.

Definition 3.3 *:* We define on T a relation of precedence noted R_c as follow: $t_1\ Rc\ t_2$ or rather t_1 precedes t_2 if the action a_1 must occur before the action a_2 (a_1 and a_2 being the actions which are the cause of e).

Proposition 3.4 : (T, R_c) is a strict order temporal framework. (T, R_c) has the discretion property , than (T, R_c) is a discrete temporal framework provided with a strict order.

f) Temporal Relationships between Events

An event can be the cause of one or more events in the future as it is often due to one or more events which proceeded in the past.

To represent this, we define the following operator which can be used to represent the effects, post and pre conditions for an action. Concept time present, past and future is represented by a relative entirety k such as:

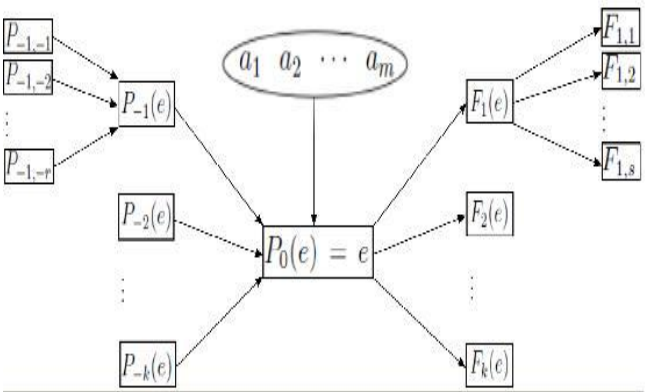


Figure 2 : Representation relationships between actions and effects/events

a) $k = 0$ represents present,
b) $k > 0$ represents the future,
c) $k < 0$ represents the past.

Definition 3.5 :

$$\otimes: \mathbb{Z}\times T \rightarrow T$$
$$(k, t) \mapsto \otimes(k, t) \equiv k \otimes t$$

- If $k = 0$, then $k \otimes t = {}_0t$ where ${}_0t = t_1 \oplus t_2 \oplus \cdots \oplus t_m$ is time-element where e occurs at the present and where m is the number of actions which are the cause of e true in ${}_0t$. We denote $e = P_0e$.
- If $k > 0$ then $k \otimes t = {}_kt$ where ${}_kt$ is time-element where the event F_ke will occur in the future and which is due to e carried in ${}_0t = t_1 \oplus t_2 \oplus \cdots \oplus t_m$
- If $k < 0$ then $k \oplus t = {}^kt$ where ${}^kt$ is time-element where the event denoted P_ke which occurred in the past and gave place to e in ${}_0t = t_1 \oplus t_2 \oplus \cdots \oplus t_m$. Here, m is the number of intervening actions so that e is true in ${}_0t$, consequently, F_ke (respectively P_ke) is true in ${}_kt$ (respectively in ${}^kt$). $|k|$ is the number of events F_ke (respectively P_ke).

The operator F_k will allow us to enumerate all effects/events that proceed in the future whereby e is the cause (ramification) and the operator P_ke will allow us to enumerate all precondition/ events which proceeded in the past and which gave place to e. The operator $\otimes$ may give us the possibility of representing the continuous evolutions of the universe for varied futures (prediction) or past (diagnostic). It may allow the representation of the actions and their effects as well as the types of reasoning which are the prediction, the explanation and planning.

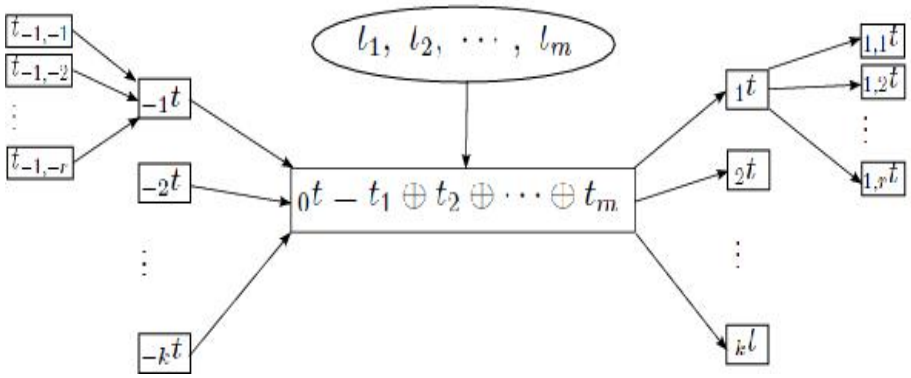


Figure 3 : Representation of temporal relationships between actions and effects/events

Figure 4: Ramified time: several past, several futures

IV. Temporal Logic L_C For Reasoning Causal Between Actions and Events/Effects

In this chapter, we propose a temporal logic to reason on the actions and events. We give its axioms and semantic.

a) Deductive System

i. Temporal Logic L_c'axioms

(i) Axioms of propositional logic [Bourbaki, 71][11].

(ii) *(a)* $F_k(A \supset_c B) = (F_kA) \supset_c (F_kB)$ where $F_k(A \supset_c B)$ is the effect/event which will occur in the future and which will take place only if $A \supset_c B$ takes place $(A \supset_c B$ is due to m actions $a_1, a_2, \cdots, a_m)$

(b) $P_k(A \supset_c B) = (P_kA) \supset_c (P_kB)$ where P_kA is an event/precondition which occurred in the past and which gave place to $(A \supset_c B)$

(c) $P_0(A \supset_c B) = (P_0A) \supset_c (P_0B)$.

The axioms (ii) : (a), (b) and (c) express the distributivity of the operators F_k, P_k and P_0 with regard to the causal implication.

ii. Temporal Logic L_c'deduction rules

The rules of deductions are :

(i) The modus ponens [Bourbaki, 71][11].

(ii) Temporal generalization: If A is a theorem, F_kA, P_kA and P_0A are equally theorems.

The theorems of L_c are by definition all the formulas deductible from the axioms by using the rules of deductions. In particular all the theorems of propositional calculus are theorems.

b) Semantic of L_c

In the semantic of propositional calculus, an assignment of values of truth V is an application, that each propositional variable associates a value of truth.

An assignment of value of truth describes a state of the world.

Definition 4.1: A valuation V on a temporal framework (T,R) is a function of set of the propositional variables in the set of the parts of T.

Definition 4.2 : A model of temporal logic is the data of a temporal framework (T,R) and a valuation V defined on this temporal framework. We note $M = (T,R,V)$.

In the case of L_c, we choose as variable propositional the actions whose effect occurs in a time-element t or actions which are the cause so that an event e is true in a time-element t.

Definition 4.3 *:* Let V_c the valuation defined on the framework temporal (T,Rc) :

$$V_c : A \longrightarrow P(T)$$

$$ai \mapsto V_c(ai) = T_i = \{t_i / a_i trueint_i\}$$

t_i is the time-element when the action a_i occurs so that the event e is true in ${}_0t = t_1 \oplus t_2 \oplus \cdots \oplus t_m$ or the effect e occurs in ${}_0t$.

The action a_i thus, occurs only once in T then $T^i = \{t_i\}$.

If T^i is empty then, a_i is not true in t_i or a was not carried out consequently e will not take place in ${}_0t$.

Definition 4.4 :

1. $V_c P_0 e = V_c(e) = V_c(a_1 \wedge \cdots \wedge a_m) =_{def} V_c(a_1) \oplus \cdots \oplus V_c(a_m) \equiv \{t_1\} \oplus \{t2\} \oplus \cdots \oplus \{t_m\} \equiv \{{}_0t\}$
2. $V_c\{\neg a_i\} = T - V_c\{a_i\} = T - T_i$
3. As e is due to the actions $a_1, a_2, \cdots, a_m$, thus, if there is k such as an action a_k does not take place, this would inevitably involve non-achievement of e (or that e will not be true in $\{{}_0t\}$ accordingly :

$V_c\{e\} = V_c\{a_1 \wedge \cdots \wedge \neg a_k \wedge \cdots \wedge a_m\} = V_c\{a_1\} \oplus \cdots \oplus V_c\{\neg a_k\} \oplus \cdots \oplus V_c\{a_m\} = T_1 \oplus \cdots \oplus \{T_k\} \oplus \cdots \oplus T_m \equiv T - V_c(e)$.

4. The effect/event e can give place to several effect/events in the future (ramification) noted $F_k e$, $k \geq 1$, and each effect/event will occur in a time-element ${}_kt$ with the following condition:

$t_i R_c {}_0t R_c k_t$ and ${}_0t = t_1 \oplus t_2 \oplus \cdots \oplus t_m$ then $V_c(F_k e) = \{{}_kt / t_i R_c {}_0t R_c k_t, {}_0t = t_1 \oplus t_2 \oplus \cdots \oplus t_m\}$.

5. the event e can be due to several events $P_k e$ which occurred in the past and each event $P_k e$ occurred in a time-element ${}_kt$ with the following condition:

$t_i R_c {}_0t R_c {}_kt {}_0t = t_1 \oplus t_2 \oplus \cdots \oplus t_m$ and therefore : $V_c(P_k e) = \{{}^k t / t_i R_c {}_0t R_c {}_k t, {}_0t = t_1 \oplus t_2 \oplus \cdots \oplus t_m\}$.

6. $V_c(A \supset_c B) = \{t / t_A R_c t_B R_c t, {}_0t = t_1 \oplus t_2 \oplus \cdots \oplus tm\}$, indeed $(A \supset_c B)$ is true in a certain time-element t pertaining to T only if A is true in one time-element t_A of T; but A true in t_A is the cause of B true in t_B, thus, to have B in t_B it is enough to have A in t_A and this will give $A \supset_c B$ true in t.

We also define the valuation in the following cases :

- Case of the effects/events require the realization of several actions at the same time. For that we define on A a relation defined as follows :

Definition 4.5 :

$$a_1 R_c a_2 \Leftrightarrow V_c(a_1) = V_c(a_2) \Leftrightarrow t_1 = t_2.$$

It will ,thus, be said that a_1 and a_2 are in relation if they occur in even time.

Proposition 4.6 *:* R_c is a relation of equivalence.

Proposition 4.7 : We have the following diagram [Mamache, 2010]:

$$\begin{array}{ccc} A & \xrightarrow{V_c} & P(T) \\ s \downarrow & & \uparrow i \\ A/R_c & \xrightarrow{\overline{V_c}} & ImV_c \end{array}$$

$\overline{V_c}(a) = V_c(a)$, $i(t_1) = \{t_1\}$ and $s(a) = a = \{a \in A / a' R_c a\}$ is the class of equivalence of a, it contains all the actions which occur at the same time as a, $ImV_c = \{s(a), a \in A\}$ is a subset of $P(T)$ and A/R_c is the set of the classes of equivalence of the elements of A, it contains the' packages' of actions or the subset of actions which are carried out at the same time in other words, the actions which occur at the same time is gathered in subsets of A in the form of classes called equivalence classes and each class is represented by an action, the time-element when this action is carried out is the time-elements of all the other actions of the class.

We can, thus, represent the set of the actions occurred at the same time by the equivalence classes of an action that is the representative of the class. We associate to this class only one time-element. This simplifies the temporal representation of actions/events.

- Case of an action which is repeated in different time-element (process). Let

Definition 4.8 :

$$\begin{array}{rccc} f: & T & \longrightarrow & A \\ & t_i & \mapsto & a_i \end{array}$$

We define on T a relation :

$$t_1 R_c t_2 \Leftrightarrow a_1 = a_2$$

it will ,thus, be said that t_1 and t_2 are in relation if the same action a occurs in t_1 and t_2.

Proposition 4.9 : R_c is a relation of equivalence.

Proposition 4.10 : We have the following diagram [Mamache, 2010][5]:

$T/R_c = \{\overline{t} / t \in T\}$, is the set of the classes of equivalence, Imf is the set of images of the elements of T, $\overline{t} = \{t_i \in T / tR_c t_i\}$ is the class of equivalence of t, it contains all the time-elements t_i where an action a produced in t and is reproduced in other time-element t_i (process).

Therefore, we represent the set of the time-elements when an action is repeated by the class of equivalence of a time-element that is the representative of the class. For this case one defines a valuation.

Definition 4.11 :

$$V_c : A \rightarrow P(T)$$
$$a \mapsto V_c(a) = \{t_i / \, a \text{ true in } t_i\}$$

- Case of competitive actions. Let a and a' two actions concurrent for the realization of an effet/event e. We have two possibilities for the choice of the actions.

Temporal choice

(i) Case where actions do not start at the same time but the agent is interested by the first achieved action,

(ii) Case where actions start at the same time but the the agent selects the action which spends less time (the least durative action),

(iii) Emergency case: the agent must choose the most urgent action.

Economic choice

(i) The agent is interested by the least expensive action in carried out independently of time,

(ii) The agent is interested by the simplest action in carried out independently of time.

Definition 4.12 : Let S a set of actions which can carry out an event e, S is a part of A. We define a relation on S:

$\forall\, a \in S, a\, R\, a' \Leftrightarrow a$ is better than a'.

An action a is the best element of S if a is better than all other actions for the realization of an event e.

Temporal choice

(i) $\forall\, a \in S, a\, \mathcal{R}\, a'$ expresses that a is the first achieved action. So, it is the action chosen by the agent,

(ii) $\forall a \in S, a\, \mathcal{R}\, a'$ expresses that a is the least durative action)

(iii) $\forall\, a \in S, a\, \mathcal{R}\, a'$ expresses a is the most urgent action.

Economic choice

(i) $\forall a \in S, a\, \mathcal{R}\, a'$ expresses that a the least expensive action in carried out independently of time,

(ii) $\forall\, a \in S, a\, \mathcal{R}\, a'$ expresses that a is the simplest action in carried out independently of time.

The corresponding valuation is defined as follows:

$$Vc : A \rightarrow P(T)$$
$$a \mapsto Vc(a) = \{ta / a \text{ true in } ta\}$$

$V_c(a) = \{ta\}$ if a' is negligible in front of a if not $V_c(a) = \emptyset$.

We can generalize this with several actions $a_1, a_2, \cdots, a_m$.

$V_c(a_i) = \{t_{ai}\}$ if a_j is negligible in front of a_i for any $j \neq i$ if not $V_c(a_i) = set$.

V. Completude

: Is Axiomatic L_c complete for the class K of the temporal framework? For that, we must show the validity : Are the theorems valid formulas ?

Theorem 5.1 (Mamache,2011) [12] Any theorem of L_c is a valid formula in the class K of the temporal framework. It should be checked that:

1) The axioms of L_c are valid formulas in K.
2) The rules of deductions preserve the validity of the formulas : if their arguments are valid, their result is true.

VI. Conclusion and Open Problem

Some basic concepts emerge in the existing actions formalisms, as causality and time. They are difficult to express in a first order language. We propose a logical formalism based on a first order language to which we add operators to represent multiples futures and multiples pasts. Furthermore, these operators allow to describe pre-conditions and effects of an action. They allow the representation of the prediction, explanation and planning.

The principal contribution of this work is the simplification of the representation of causal and temporal relationships between actions and their effects as well as the causal and temporal relationships between actions and events. We used the classes of equivalence to represent the execution time of a process and the execution time of competitive actions. We propose a new ontology for causal and temporal representation of actions/events. The ontology used in our formalism consists of facts, events, process, causality, action and planning.

Although this work is located in axis of theoretical study of knowledge reasoning, we can hope that this study will be used as a basis on which theories of action can be established. It can be prolonged in several directions.

- A track which appears very important consists in representing temporal relationships of the causes of events if these causes are complex actions/events. We plan a matrix representation to enrich our formalism.
- Inspired of action modeling formalisms, more precisely of action theory and Allen's time [Allen, 84][1], Galton [Galton, 2009][15]has combined a space theory with a temporal theory. The primitive entities of Galton are moments and intervals but he does not consider the cases where the regions are separate in the future and the past.

Inspired by our ontology, we envisage a new ontology to represent space-time relationships between objects and regions where the events will be the changes caused by the various positions of the objects. Our formalism could be used to facilitate space-time

representation of objects positions. This logic allows to study the evolution of the relative positions between entities during time.

- F.Baader and al [Baader, 2005][16]propose an action formalism based on description logics (Dls). H.Strass and M.Thielscher [11] study the integration of two prominent fields of logic- based AI: action formalisms and non-monotonic reasoning.
- H. Liu[Liu, 2010][18] have investigated updates of ABoxes in DLs and analyzed their computational behavior. The main motivation for this en-deavor is to establish the theoretical foundations of progression in action theory based on DLs and to provide support for reasoning about action in DLs. Within the framework of integration description logics in action formalism, we envisage to integrating description logics in our temporal formalism.
- The information extraction (IE) is an important subject of research in Natural Languages Automatic Processing . The analysis of named entities (EN)[Ehrmann, 2008][14] is generally focused on the traditional concepts of place, organization, person or dates. The events are rarely considered, but they have a great importance for the usual applications as information search. Our formalism can be used to develop information extraction system of extraction of event type.
- Another study way would be the exploitation of temporality in biographical information extraction . The temporal reference marks of biographical information allow to replace a fact in its context and to order it compared with other events by using our operators F_K and P_K. A good exploitation of our approach will certainly make it possible to obtain a functional and satisfactory solution with the problems encountered within the framework of the extraction and the information management.
- In medical applications, our formalism can to be used to describe states of the world, such as data of patients. In this context, the actions can be used to represent diagnostic and therapeutic of the processing of patient treatment.
- Several experimental works will certainly make it possible to enrich this work in particular, like implementing an interface to represent expressions of temporal actions type and events temporal type based on our formalism. This work would allow to describe several applications and to compare them with other formalisms.

References Références Referencias

1. J.Allen, "Towards a general theory of actions and time", *Artificial Intelligence*, 23, (1984), pp.123-154
2. J. F. Allen and G. Ferguson, " Actions and Events in Interval Temporal Logic", *J. Logic and Computation*, 4, 5, (1994)
3. D.v McDermott, "A temporal logic for reasoning about processes and plans", *Process and Cognitive Science*, 6, (1982), pp.101-155.
4. D.Kayser, A.Mokhtari, " Time in a Causal Theory", *Annals of Mathematics and Artificial Intelligence*, 22(1-2),117-138 (1998)
5. F.Mamache, "Events, Actions and Temporal logic", *International Journal of Open Problems in Computer Science and Mathematics*, Vol.3, No.3 (2010)
6. B. Knight, J.My, T. Peng, "Reasoning about Changes over Time: Actions, Events, and their Effects", *Proc. Formalization of Commonsense Reasoning*, (1998), pp.183-197.
7. B. Knight, J. My, T. Peng, "Representing Temporal Relationships Between Events and Their Effects", *Proceedings of the International Fourth Temporal Workshop one Representation and Reasoning, IEEE Computer Society Press*, (1997), pp.148-152.
8. A.Pinto, "Temporal reasoning in the situtin calculus", PH D. thesis, *Computer Science, University of Toronto*, (1994)
9. H. Brstougeff and G. Ligozat," Outils logiques pour le traitement du temps de la linguistique à l'intelligence artificielle", *Masson, Paris*, (1989)
10. S.C Kleene, "Mathematical Logic",*Collection U*, (1971).
11. N.Bourbaki, "Théorie des ensembles, éléments de mathématiques", *Hermann, Paris*, (1971).
12. F.Mamache, "A Temporal Logic for Reasoning about Actions", *International Journal of Open Problems in Computer Science and Mathematics*, Vol.4, No.1 (2011)
13. F. Even, Extraction d'Information et modalisation de connaissances à partir de Notes de Communication Orale PH D. thesis, *Université de Nantes*, (2005)
14. M. EHRMANN, Les entités Nommées, de la linguistique au TAL : Statut théorique et méthodes de désambiguisation PH D. thesis, *Université de Paris*, (2008).
15. A.Galton, "Spatial and Temporal Knowledge representation", *Journal of Earth Science Informatics* 2(3)169-187 (2009)
16. F. Baader, C. Lutz, M. Milicic, U. Sattler, and F. Wolter, "Integrating description logics and action formalisms: First results", In *Proceedings of the Twentieth National Conference on Artificial Intelligence* (AAAI05),*Pittsburgh, PA, USA,*(2005).
17. H. Strass, M. Thielscher, "On Defaults in Action Theories", In *Proceedings of the 32nd German Annual Conference on Artificial Intelligence (KI'09), Paderborn, Germany, Springer-Verlag Berlin Heidelberg* , (2009), pp.298-305.
18. H.Liu, *Computing Updates in Description Logics*, PH D. thesis, Dresden University of Technology, Germany, (2010).

Global Journals Inc. (US) Guidelines Handbook 2012

www.GlobalJournals.org

FELLOW OF ASSOCIATION OF RESEARCH SOCIETY IN COMPUTING (FARSC)

- 'FARSC' title will be awarded to the person after approval of Editor-in-Chief and Editorial Board. The title 'FARSC" can be added to name in the following manner. eg. **Dr. John E. Hall, Ph.D., FARSC or William Walldroff Ph. D., M.S., FARSC**
- Being FARSC is a respectful honor. It authenticates your research activities. After becoming FARSC, you can use 'FARSC' title as you use your degree in suffix of your name. This will definitely will enhance and add up your name. You can use it on your Career Counseling Materials/CV/Resume/Visiting Card/Name Plate etc.
- 60% Discount will be provided to FARSC members for publishing research papers in Global Journals Inc., if our Editorial Board and Peer Reviewers accept the paper. For the life time, if you are author/co-author of any paper bill sent to you will automatically be discounted one by 60%
- FARSC will be given a renowned, secure, free professional email address with 100 GB of space eg.johnhall@globaljournals.org. You will be facilitated with Webmail, Spam Assassin, Email Forwarders, Auto-Responders, Email Delivery Route tracing, etc.
- FARSC member is eligible to become paid peer reviewer at Global Journals Inc. to earn up to 15% of realized author charges taken from author of respective paper. After reviewing 5 or more papers you can request to transfer the amount to your bank account or to your PayPal account.
- Eg. If we had taken 420 USD from author, we can send 63 USD to your account.
- FARSC member can apply for free approval, grading and certification of some of their Educational and Institutional Degrees from Global Journals Inc. (US) and Open Association of Research, Society U.S.A.
- After you are FARSC. You can send us scanned copy of all of your documents. We will verify, grade and certify them within a month. It will be based on your academic records, quality of research papers published by you, and 50 more criteria. This is beneficial for your job interviews as recruiting organization need not just rely on you for authenticity and your unknown qualities, you would have authentic ranks of all of your documents. Our scale is unique worldwide.
- FARSC member can proceed to get benefits of free research podcasting in Global Research Radio with their research documents, slides and online movies.
- After your publication anywhere in the world, you can upload you research paper with your recorded voice or you can use our professional RJs to record your paper their voice. We can also stream your conference videos and display your slides online.
- FARSC will be eligible for free application of Standardization of their Researches by Open Scientific Standards. Standardization is next step and level after publishing in a journal. A team of research and professional will work with you to take your research to its next level, which is worldwide open standardization.

- FARSC is eligible to earn from their researches: While publishing his paper with Global Journals Inc. (US), FARSC can decide whether he/she would like to publish his/her research in closed manner. When readers will buy that individual research paper for reading, 80% of its earning by Global Journals Inc. (US) will be transferred to FARSC member's bank account after certain threshold balance. There is no time limit for collection. FARSC member can decide its price and we can help in decision.

MEMBER OF ASSOCIATION OF RESEARCH SOCIETY IN COMPUTING (MARSC)

- 'MARSC' title will be awarded to the person after approval of Editor-in-Chief and Editorial Board. The title 'MARSC" can be added to name in the following manner. eg. Dr. John E. Hall, Ph.D., MARSC or William Walldroff Ph. D., M.S., MARSC
- Being MARSC is a respectful honor. It authenticates your research activities. After becoming MARSC, you can use 'MARSC' title as you use your degree in suffix of your name. This will definitely will enhance and add up your name. You can use it on your Career Counseling Materials/CV/Resume/Visiting Card/Name Plate etc.
- 40% Discount will be provided to MARSC members for publishing research papers in Global Journals Inc., if our Editorial Board and Peer Reviewers accept the paper. For the life time, if you are author/co-author of any paper bill sent to you will automatically be discounted one by 60%
- MARSC will be given a renowned, secure, free professional email address with 30 GB of space eg.johnhall@globaljournals.org. You will be facilitated with Webmail, Spam Assassin, Email Forwarders, Auto-Responders, Email Delivery Route tracing, etc.
- MARSC member is eligible to become paid peer reviewer at Global Journals Inc. to earn up to 10% of realized author charges taken from author of respective paper. After reviewing 5 or more papers you can request to transfer the amount to your bank account or to your PayPal account.
- MARSC member can apply for free approval, grading and certification of some of their Educational and Institutional Degrees from Global Journals Inc. (US) and Open Association of Research, Society U.S.A.
- MARSC is eligible to earn from their researches: While publishing his paper with Global Journals Inc. (US), MARSC can decide whether he/she would like to publish his/her research in closed manner. When readers will buy that individual research paper for reading, 40% of its earning by Global Journals Inc. (US) will be transferred to MARSC member's bank account after certain threshold balance. There is no time limit for collection. MARSC member can decide its price and we can help in decision.

AUXILIARY MEMBERSHIPS

ANNUAL MEMBER

- Annual Member will be authorized to receive e-Journal GJMBR for one year (subscription for one year).
- The member will be allotted free 1 GB Web-space along with subDomain to contribute and participate in our activities.
- A professional email address will be allotted free 500 MB email space.

PAPER PUBLICATION

- The members can publish paper once. The paper will be sent to two-peer reviewer. The paper will be published after the acceptance of peer reviewers and Editorial Board.

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 .

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.

Administration Rules Listed Before Submitting Your Research Paper to Global Journals Inc. (US)

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

K

M

N

O

P

R

S

T

U

V

W

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