Global Journal of Computer Science and Technology

discovering thoughts and inventing future

Volume 10 Issue 4 Version 1.0 | Online ISSN: 0975-4172 | Print ISSN: 0975-4350

21 Technology Changing Ideas

highlights

- SHA-256 Hashing Cryptographic Module
- Embryonic Machine
- Virtual Backbone Routing
- Polynomial Regression Modules

June 2010

ENG

Global Journal of Computer Science and Technology

Global Journal of Computer Science and Technology

Volume 10 Issue 10 (Ver. 1.0)

Global Academy of Research and Development

License No.: 42125/022010/1186
Registration No.: 430374
Import-Export Code: 1109007027

Global Academy of Research and Development

Publisher's correspondence office

Global Journals Inc., Headquarters Corporate Office, United States

Offset Typesetting

Global Journals Inc., City Center Office, United States

Packaging & Continental Dispatching

Global Journals Inc., 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) & 500 USD (Color)

Editorial Board Members

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

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

Dr. Fotini Labropulu
Mathematics - Luther College
University of 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 School Ph.D. (WHU Koblenz)
MBA/BBA (University of Saarbrücken)

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

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

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

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

Dr. Pina C. Sanelli
Associate Professor of Public Health
Weill Cornell Medical College
Associate Attending Radiologist
New York-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, Taiwan University Chair Professor
Institute of Environmental Engineering, National Chiao Tung University, Hsin-chu, Taiwan.Ph.D., MS The University of Chicago, Geophysical Sciences
BS National Taiwan University, Atmospheric Sciences
Associate Professor of Radiology

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

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

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

Chief Author

Dr. R.K. Dixit (HON.)
M.Sc., Ph.D., FICCT
Chief Author, India
Email: authorind@computerresearch.org

Dean & Editor-in-Chief (HON.)

Vivek Dubey(HON.)
MS (Industrial Engineering),
MS (Mechanical Engineering)
University of Wisconsin
FICCT
Editor-in-Chief, USA
editorusa@computerresearch.org

Er. Suyog Dixit
BE (HONS. in Computer Science), FICCT
SAP Certified Consultant
Technical Dean, India
Website: www.suyogdixit.com
Email:suyog@suyogdixit.com,
dean@computerresearch.org

Sangita Dixit
M.Sc., FICCT
Dean and Publisher, India
deanind@computerresearch.org

Contents of the Volume

From the Chief Author's Desk

We see a drastic momentum everywhere in all fields now a day. Which in turns, say a lot to everyone to excel with all possible way. The need of the hour is to pick the right key at the right time with all extras. Citing the computer versions, any automobile models, infrastructures, etc. It is not the result of any preplanning but the implementations of planning.

With these, we are constantly seeking to establish more formal links with researchers, scientists, engineers, specialists, technical experts, etc., associations, or other entities, particularly those who are active in the field of research, articles, research paper, etc. by inviting them to become affiliated with the Global Journals.

This Global Journal is like a banyan tree whose branches are many and each branch acts like a strong root itself.

Intentions are very clear to do best in all possible way with all care.

Dr. R. K. Dixit
Chief Author
chiefauthor@globaljournals.org

Improper Data Collection Mechanisms, an Important Cause for Erroneous Corporate Metrics

Alexandru Petchesi

GJCST Classification
H.2.8 I.2.4

Abstract- **This paper intends to highlight one of the very important but most often overlooked aspects related to the challenges of the customization of information systems due to the lack of repeatability and reproducibility during data collection.**

Keywords- Knowledge Management, Data Validation, Repeatability and reproducibility of data collection, Corporate metrics.

I. INTRODUCTION

Many companies in the 21st century are monitoring their regular activities through performance metrics. To calculate performance metrics data needs to be collected in a well defined way and stored for analysis purposes. To accomplish this goal many companies invest significant amounts of money into information systems such as Enterprise Resource Planning and Manufacturing Execution Systems to collect and report such data (Fig.1).

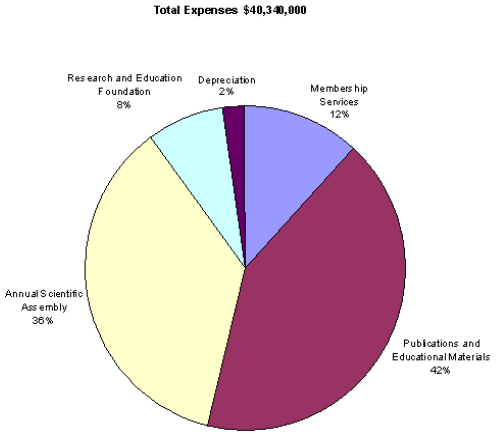


Fig 1: Corporate metrics: Pie chart of expenses

The strategy that many companies use to implement their information highway is through the acquisition of off-the-shelf solutions which are then customized to the needs of the company. As currently there is no one software solution that can provide all information services needed by a company, the solution to build an information highway adopted by many companies is to purchase best of breed solutions and then integrate them. These integrations presented and will present quite a lot of challenges to companies due to the large variety of technologies used to implement them. This paper intends to highlight one of the aspects related to the challenges of the customization of information systems due to the lack of repeatability and reproducibility during data collection.

About-Alexandru Petchesi
Master in I.T. Management, University of Oradea, Oradea, Romania
H.BSc. Computer Science, McMaster University, Hamilton, CanadaE-Mail: petcheai@yahoo.com

II. THE PROBLEM

What can go wrong during data collection process that can affect the quality and quantity of data we are collecting and therefore the reports we are generating from our information systems? I would like to present in this paper one of the major risk factors that has an impact on the data collected by an information system, the repeatability and reproducibility of the data collection process.The problem will be exemplified with a very simple case, for a shop floor control system. Imagine working in a manufacturing company that produces a certain product. One of the very important metrics related to manufacturing a product is the quality of the product which is measured in most companies through metrics such as first pass or rolled throughput yield. To calculate metrics such as the ones mentioned above, companies need to collect information on the products they manufacture such as the number of products with defects. An important characteristic of the data is related to its granularity , mainly related to the categories of defects that can be identified on a product. The data collection process of such data is done in many companies through operators which need to visually identify the cause of failure, then pick their data manually from a list of options offered to them by the software. Data collected in this manner lead to reports such as the Pareto chart presented below that gives decission makers within a company the information needed to identify the route causes of problems and take the necessary actions based on them. Therefore the accuracy of such a data collection process is very important as a report as the one presented below gives people in a company an image of the realities within the production process from within a company. If data is —distorted" the image provided through the reporting mechanism is also distorted and does not reflect the realities from the factory.

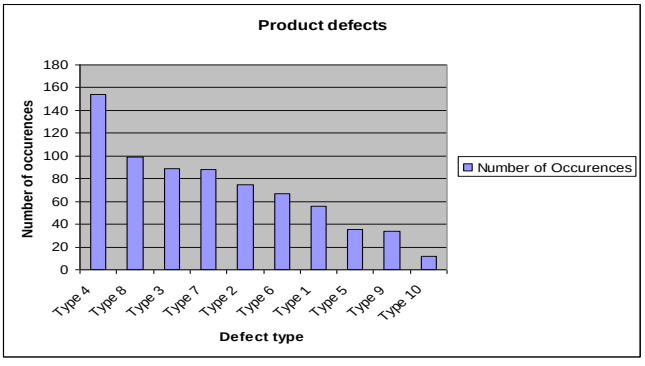


Fig.2 Pareto chart of the product defects from a manufacturing company

The reports as the one presented in Fig.2 provide companies images of the realities from within the factory and give them clues on the area of the process where they need to act upon to start improvement projects.People in information technology are very familiar with the —Garbage in Garbage Out" principle. To reduce or even completely eliminate this problem from a software application, the information technology community has developed defensive programming techniques. One of the best practices of data collection tells us that, in order to assure that the data we collect from the end users is right it is preferred to employ in a in the graphical user interface of a software application, pre-defined selection mechanisms such as combo boxes, selection lists etc. These allow the end-users to easily perform single or multiple selections of values from a well defined set.The data set for such a list is usually defined by the subject matter expert working on the business side with IT experts in charge with the customization of the tool. During the lifetime of the software product, employees using the software will use such a combo box to pick the proper values and submit them into the database for storage. When we are inputting data into an information system, the IT best practices are telling us, we need to make sure that we avoid the garbage in garbage out principle. The major effect of the pronciple above is that once the data collected is —contaminated" in our storage it will affect all the information systems from within our architecture that use this data source as the master. The result is that erroneous information will spread all across the company and this information can cost us significant amount of data due to spending the company might make as a result of the reports provided (Fig.3).

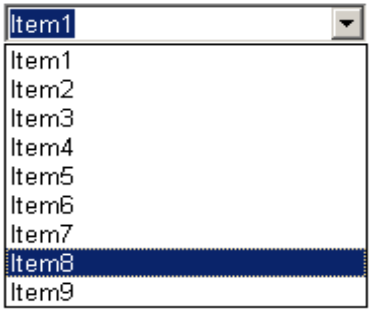


Fig 3. A combo box

What can go wrong during such a data collection mechanism that can affect the quality and quantity of data we are collecting? Everything seems to be properly set up from an IT prespective, but the employees of the company using the reports are sometimes complaining about discrepancies between the realities they are aware of and the data from the reports. Many of them become quickly frustrated and start loosing the trust on the reports provided to them many times by expensive software tools with a steep learning curve.The experiment presented below will identify an overlooked way of erroneous data entering an information system due to the lack of repeatability and reproducibility of the data collection process.

III. THE EXPERIMENT

We are going to illustrate the repeatability and reproducibility issues of data collection through an example from the electronic manufacturing industry. The experiment was conducted a long time ago and the purpose of it in this paper is for exemplification only. The experiment will present the Gage R&R methodology from Six Sigma, an important statistical tool that can allow us to determine the repeatability and reproducibility of a data input process.

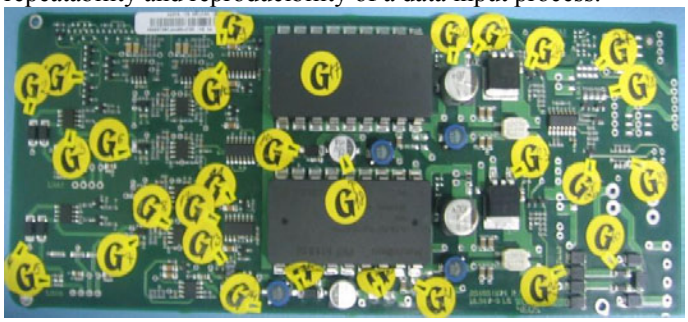


Fig 4. A printed circuit board with 30 different marked locations on it

Sample #	Defect Definition
1	Missing
2	Misoriented
3	Ok
4	Insufficient Solder
5	Missing
6	Ok
7	Misoriented
8	Misoriented
9	Ok
10	Misoriented
11	Ok
12	Solder Bridging
13	Misoriented
14	Additional Component
15	Excess Solder
16	Misoriented
17	Misoriented
18	Misoriented
19	Additional Component
20	Ok
21	Damaged
22	Additional Component
23	Wrong Part Number
24	Misplaced
25	Misoriented
26	Ok
27	Ok
28	Ok
29	Misplaced
30	Ok

Table 1 The list of issues for each location on the board (the standard)

A printed circuit board (Fig. 4) was used and marked with 30 locations, some locations marked had defects some did not have any defects, to identify the accuracy of the data collection mechanism. Three operators were selected randomly to determine how close their selection of data from a particular set was to the standard (Table. 4).

The three operators selected were presented with a set of allowable values and they were asked to pick defects from a list of standard defects providedby in information system. This data selection mechanism was used by them already in their daily activities through an information system, using data provied by a combo box, where they needed to select one value from a list. Their answers were collected in a spreadsheet and in case their answer matched the standard defined by the expert a PASS was introduced in the Gage R&R tool and a FAIL was introduced in case their selection did not match the standard. (Fig.5).

The experiment was repeated a week later with the same operators on the same printed circuit board without informing them about the fact that it was the same product. The data collected from the second session was introduced in a similar way in the Gage R&R tool, as seen below. The spreadsheet then calculated for us the differences between what each operator‘s option and the standard defined by the expert providing us very valuable information on the data identification and selection mechanism.

Using the Gage R&R method we looked at the consistency of the data selection mechanism for each individual, between individuals and against the standard. The conclusion we drew were pretty interesting!

SCORING REPORT

DATE: 26.09.2006
NAME: Expert
PRODUCT: Defect identification
BUSINESS: VISUAL INSPECTION

Attribute Legend[5] (used in computations)
1 PASS
2 FAIL

Known Population		Operator 1		Operator 2		Operator 3		All operators agree within and between each Other	All Operators agree with standard
Sample #	Attribute	Try #1	Try #2	Try #1	Try #2	Try #1	Try #2	Y/N Agree	Y/N Agree
1	PASS	PASS	PASS	PASS	PASS	PASS	PASS	Y	Y
2	PASS	PASS	PASS	PASS	PASS	PASS	PASS	Y	Y
3	PASS	PASS	PASS	PASS	PASS	PASS	PASS	Y	Y
4	PASS	FAIL	FAIL	PASS	PASS	FAIL	FAIL	N	N
5	PASS	PASS	PASS	PASS	PASS	PASS	PASS	Y	Y
6	PASS	FAIL	FAIL	PASS	PASS	FAIL	FAIL	N	N
7	PASS	FAIL	PASS	PASS	PASS	PASS	FAIL	N	N
8	PASS	PASS	PASS	PASS	PASS	PASS	PASS	Y	Y
9	PASS	FAIL	PASS	PASS	PASS	FAIL	FAIL	N	N
10	PASS	PASS	PASS	PASS	PASS	PASS	PASS	Y	Y
11	PASS	PASS	PASS	PASS	PASS	PASS	PASS	Y	Y
12	PASS	PASS	PASS	PASS	PASS	PASS	PASS	Y	Y
13	PASS	PASS	PASS	PASS	PASS	PASS	PASS	Y	Y
14	PASS	PASS	PASS	PASS	PASS	PASS	PASS	Y	Y
15	PASS	PASS	PASS	PASS	PASS	PASS	PASS	Y	Y
16	PASS	PASS	PASS	PASS	PASS	PASS	PASS	Y	Y
17	PASS	PASS	PASS	PASS	PASS	PASS	PASS	Y	Y
18	PASS	PASS	PASS	FAIL	FAIL	FAIL	FAIL	N	N
19	PASS	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL	Y	N
20	PASS	PASS	PASS	PASS	PASS	FAIL	PASS	N	N
21	PASS	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL	Y	N
22	PASS	PASS	FAIL	PASS	PASS	FAIL	PASS	N	N
23	PASS	FAIL	FAIL	PASS	PASS	PASS	PASS	N	N
24	PASS	PASS	PASS	PASS	FAIL	FAIL	PASS	N	N
25	PASS	PASS	PASS	PASS	FAIL	PASS	PASS	N	N
26	PASS	PASS	PASS	PASS	PASS	PASS	FAIL	N	N
27	PASS	PASS	PASS	PASS	PASS	PASS	PASS	Y	Y
28	PASS	PASS	PASS	PASS	PASS	PASS	FAIL	N	N
29	PASS	PASS	PASS	PASS	PASS	FAIL	PASS	N	N
30	PASS	PASS	PASS	PASS	PASS	PASS	PASS	Y	Y

Fig 5. Gage R&R with data collected from the three operators

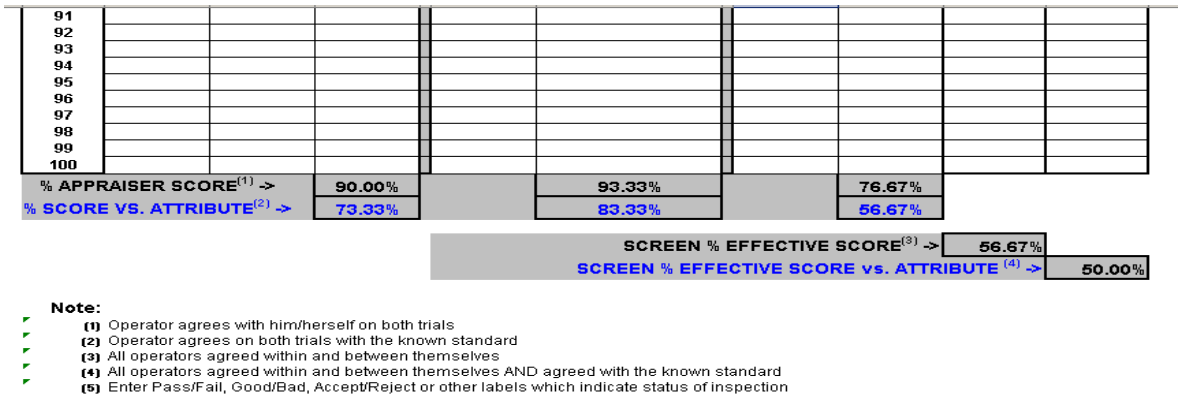


Fig 6. Gage R&R Statistics

The statistical analysis of the repeatability and reproducibility of the data selection process are shown in Fig.6.

IV. Conclusions

As seen in the results above (Fig. 6) there are significant discrepancies for all 4 categories tested. The report tells us that the values picked from the list by the operators and entered in the information system and the realities as defined by the standard are significantly different. If this data would have been entered into an information system the reports generated from the data entered would have been very different from the realities from the factory and actions might have been taken in the wrong direction by the team using reports based on the data. Therefore, it is important that any data entry process which collects data introduced by human operators based on non-numeric criteria must be validated on regular basis for the repetability and reproductibility. Without this validation the money invested in information systems will not provide the value adds they were purchased for and can even produce significant financial losses to companies due to erronous reporting.

A Multicast Protocol For Content-Based Publish-Subscribe Systems

GJCST Classification
C.2.2 , H.3.5

Yashpal Singh[1], Vishal Nagar[2], D.C.Dhubkarya1

Abstract- The publish/subscribe (or pub/sub) paradigm is a simple and easy to use model for interconnecting applications in a distributed environment. Many existing pub/sub systems are based on pre-defined subjects, and hence are able to exploit multicast technologies to provide scalability and availability. An emerging alternative to subject-based systems, known as content-based systems, allow information consumers to request events based on the content of published messages. This model is considerably more flexible than subject-based pub/sub, however it was previously not known how to efficiently multicast published messages to interested content-based subscribers within a network of broker (or router) machines. In this paper, we develop and evaluate a novel and efficient technique for multicasting within a network of brokers in a content-based subscription system, thereby showing that content-based pub/sub can be deployed in large or geographically distributed settings.

Keywords- Multicast , Publishers, Subscribers, information spaces, OMG, content-based routing

I. Introduction

The publish/subscribe paradigm is a simple, easy to use and efficient to implement paradigm for interconnecting applications in a distributed environment. Pub/sub based middleware is currently being applied for application integration in many domains including financial, process automation, transportation, and mergers and acquisitions. Pub/sub systems contain information providers, who publish events to the system, and information consumers, who subscribe to particular categories of events within the system. The system ensures the timely delivery of published events to all interested subscribers. A pub/sub system also typically contains message brokers that are responsible for routing messages between publishers and subscribers.

The earliest pub/sub systems were subject-based. In these systems, each unit of information (which we will call an event) is classified as belonging to one of a fixed set of subjects (also known as groups, channels, or topics). Publishers are required to label each event with a subject; consumers subscribe to all the events within a particular subject. For example a subject-based pub/sub system for stock trading may define a group for each stock issue; publishers may post information to the appropriate group, and subscribers may subscribe to information regarding any issue. In the past decade, systems supporting this paradigm have matured significantly resulting in several academic and industrial strength solutions [4][10][12][13][15]. A similar approach has been adopted by the OMG for CORBA event channels [11].An emerging alternative to subject-based systems is content-based subscription systems [6][14]. These systems support a number of information spaces, each associated with an event schema defining the type of information contained in each event. Our stock trade example (shown in Figure 1) may be defined as a single information space with an event schema defined as the tuple [issue: string, price: dollar, volume: integer]. A content-based subscription is a predicate against the event schema of an information space, such as (issue="IBM" & price < 120 & volume > 1000) in our example.With content-based pub/sub, subscribers have the added flexibility of choosing filtering criteria along as many dimensions as event attributes, without requiring pre-definition of subjects. In our stock trading example, the subject-based subscriber is forced to select trades by issue name. In contrast, the content-based subscriber is free to use an orthogonal criterion, such as volume, or indeed a collection of criteria, such as issue, price and volume. Furthermore, content-based pub/sub removes the administrative overhead of defining and maintaining a large number of groups, thereby making the system easier to manage. Finally, content-based pub/sub is more general in that it can be used to implement subject-based pub/sub, while the reverse is not true. While content-based pub/sub is the more powerful paradigm, efficient and scalable implementations of such systems have previously not been developed.

In order to efficiently implement a content-based pub/sub system, two key problems must be solved:

i. The problem of efficiently matching an event against a large number of subscribers on a single message broker.
ii. The problem of efficiently multicasting events within a network of message brokers. This problem becomes crucial in two settings: 1) when the pub/sub system is geographically distributed and message brokers are connected via a relatively low speed WAN (compared to high-speed LANs), and 2) when the pub/sub system has to scale to support a large number of publishers, subscribers and events. In both cases it becomes crucial to limit the distribution of a published event to

Yashpal Singh[1], Vishal Nagar[2], D.C.Dhubkarya[1]
yash_biet@yahoo.com, vishal_biet@yahoo.com, dcd3580@yahoo.com
1.BIET Jhansi, 2. CSE Jhansi , India

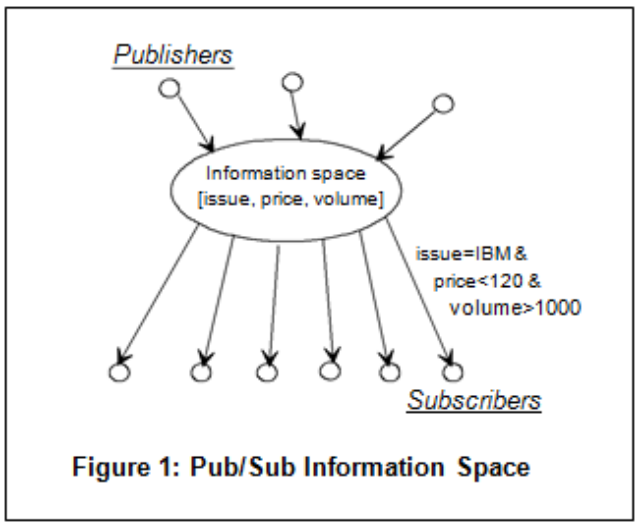


Figure 1: Pub/Sub Information Space

only those brokers that have subscribers interested in that event.One of the strengths of subject-based pub/sub systems is that both problems are trivial to solve: the matching problem is solved using a mere table lookup; the multicast problem is solved by defining a multicast group per subject, and multicasting each event to the appropriate multicast group. For content-based pub/sub systems, however, previous literature does not contain solutions to either problem, matching or multicasting. In this paper we present the first efficient solution to the multicast problem for content-based pub/sub. In a companion paper [2] we present an efficient soution to the matching problem for these systems.There are two straightforward approaches to solving the multicasting problem for content-based systems: (1) in the *match-first* approach, the event is first matched against all subscriptions, thus generating a destination list and the event is then routed to all entries on this list; and (2) in the *flooding* approach, the message is broadcast or flooded to all destinations using standard multicast technology and unwanted messages are filtered out at these destinations. The match-first approach works well in small systems, but in a large system with thousands of potential destinations, the increase in message size makes the approach impractical. Further, with this approach we may have multiple copies of the same message going over the same network link on its way to multiple remote subscribers. The flooding approach suffers when, in a large system, only a small percentage of clients want any single message. Furthermore, the flooding technique cannot exploit *locality* of information requests, i.e., when clients in a single geographic area are, for many applications, likely to have similar requests for data.

The central contribution of this paper is a new protocol for content-based routing, an efficient solution to the multicast problem for content-based pub/sub systems. With this protocol, called link matching, each broker partially matches events against subscribers at each hop in the network of brokers to determine which brokers to send the message. Further, each broker forwards messages to its subscribers based on their subscriptions. The disadvantages of the match-first approach are avoided since no additional information is appended to the message headers. Further, at most one copy of a message is sent on each link. The disadvantages of the flooding approach are avoided as the message is only sent to brokers and clients needing the message, thus exploiting locality. We illustrate, using a network simulator, that flooding overloads the network at significantly lower publish rates than link matching. We also describe our implementation of a distributed Java based prototype of content-based pub/sub brokers.

The remainder of this paper is organized as follows. In section 2, we present a solution to the matching problem (i.e., the case when the network consists of a single broker). In section 3, we discuss how to extend the solution to the matching problem into a solution to the content-based routing problem in a multi-broker network. In section 4, we evaluate the performance of this approach and compare it to the flooding approach.

II. THE MATCHING ALGORITHM

This section summarizes a non-distributed algorithm for matching an event against a set of subscriptions, and returning the subset of subscriptions that are satisfied by the event. (A more detailed presentation of matching along with experimental and analytic measures of performance are the subject of our companion paper [2].) This matching algorithm is the basis of our distributed multicast protocol, presented in the following section.

Our approach to matching is based on sorting and organizing the subscriptions into a parallel search tree (or PST) data structure, in which each subscription corresponds to a path from the root to a leaf. The matching operation is performed by following all those paths from the root to the leaves that are satisfied by the event. Intuitively, this data structure yields a scaleable algorithm because it exploits the commonality between subscriptions as shared prefixes of paths from root to leaf.

Figure 2 shows an example of a matching tree for an event schema consisting of five attributes a1 through a5. These attributes could represent, for example, the stock issue, price, or volume attributes mentioned above. The root of the tree corresponds to a test of the value of attribute a1, the nodes at the next level correspond to a test of attribute a2, etc. The branches are labeled with the values of the attributes being tested. In the example, we only show equality tests (although range tests are also possible), so the right branch of the root represents the test a1 = 1. The left branch of the root, with label *, means that the subscriptions along that branch do not care about the value of the attribute. Each leaf is labeled with the identifiers of all the subscribers wishing to receive events matching the predicate, i.e., all the tests from the root to the leaf. For example, in Figure 2, the rightmost leaf corresponds to a subscription whose predicate is (a_1=1 & a_2=2 & a_3=3 & a_5=3). Since a_4 does not appear in this subscription, it is represented by a label * in the PST.

Given this tree representation of subscriptions, the matching algorithm proceeds as follows. We begin at the root, with current attribute a_1. At any non-leaf node in the tree, we find

the value v_j of the current attribute a_j. We then traverse any of the following edges that apply: (1) the edge labeled v_j if there is one, and (2) the edge labeled * if there is one. This may lead to either 0, 1, or 2 successor nodes (or more in the general case where the tests are not all strict equalities). We initiate parallel subsearches at each successor node. When any of the parallel subsearches reaches a leaf, all subscriptions at that leaf are added to the list of matched subscriptions. For example, running the matching algorithm with the matching tree of Figure 2 and the event a = <1, 2, 3, 1, 2> will visit all the nodes marked with dark circles and will match four subscription predicates, corresponding to the dark circles at leaf nodes.

The way in which attributes are ordered from root to leaf in the PST can be arbitrary. In our experience, however, performance seems to be better if the attributes near the root are chosen to have the fewest number of subscriptions labeled with a *.

In the companion paper [2], we have analytically shown that the cost of matching using the above algorithm increases *less* than linearly as the number of subscriptions increase.

A. Optimizations

A number of optimizations may be applied to the parallel search tree to decrease matching time -- these optimizations are explained fully in [2].

Factoring- Some search steps can be avoided, at the cost of increased space, by factoring out certain attributes. That is, certain attributes --- preferably those for which the subscriptions rarely contain ―don't care" tests --- are selected as indices. A separate subtree is built for each possible value (or for ranges, each distinguished value range) of the index attributes.

Trivial Test Elimination- Nodes with a single child which is reached by a *-branch may be eliminated.

Delayed Branching- Following *-branches may be delayed until after a set of tests have been applied. This optimization prunes paths from that *-branch which are inconsistent with the tests.

It is worth noting that, under certain circumstances, after applying optimizations, the parallel search tree will no longer be a tree but instead a directed acyclic graph.

III. THE LINK MATCHING ALGORITHM

The previous section described a non-distributed algorithm for matching events to subscriptions. This section presents the central contribution of this paper -- an extended matching algorithm for a network of brokers, publishers, and subscribers (as shown in Figure 3). The problem, in this case, is to efficiently deliver an event from a publisher to all distributed subscribers interested in the event.

One straightforward solution to this problem is to perform the matching algorithm of the previous section at the broker nearest to the publisher, producing a destination list consisting of the matched subscribers. This destination list may be undesirably long in a large network with thousands of subscribers, and it may be infeasible to transmit and process large messages containing long destination lists throughout the network.

Link matching is our strategy for multicasting events without using destination lists. After receiving an event, each broker receiving an event performs just enough matching steps to determine which of its *neighbors* should receive it. As shown in Figure 3, neighbors may be brokers

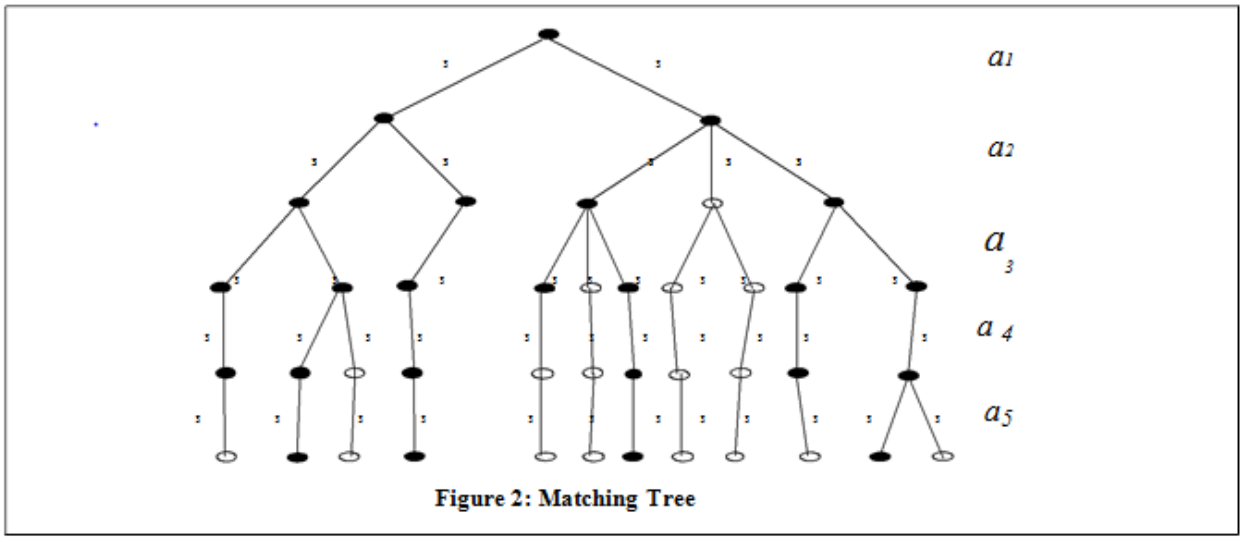


Figure 2: Matching Tree

or clients (this figure shows a spanning tree derived from the actual non-tree broker network). That is, each broker, rather than determining which subset of all subscribers is to receive the event, instead computes which subset of its neighbors is to receive the event, i.e., it determines those links along which it should transmit the message. Intuitively, this approach should be more efficient because the number of links out of a broker is typically much less than the total number of subscribers in the system.

To perform link matching, we use the parallel search tree (PST) structure of the previous section, where each path

from root to leaf represents a subscription. We augment the PST with vectors of trits, where the value of each trit is either —Yes," (Y) —No," (N) or —Maybe" (M).

We begin by annotating leaf nodes in the PST with a trit vector of size equal to the number of links out of that broker. For each link out of a broker, a position in a trit vector determines whether to send matched events down that link, based on whether there exists a subscription reachable via that link. Leaf annotations are then propagated to non-leaf nodes in a bottom-up manner. A "Yes" in a trit annotation means that (based on the tests performed so far) the event will be matched by some subscriber that is best reached by sending the message along the given link; —No" means that the event will definitely not be matched by any subscriber along that link; and —Maybe" means that further searching must take place to determine whether or not there is such a subscriber. Annotations are described in more detail below.

The link matching algorithm consists of the following three steps. First, at each broker, the parallel search tree is annotated with a trit vector encoding link routing information for the subscriptions in the broker network. Second, an initialization mask of trits must be computed at each broker for each spanning tree used for message routing. (Collectively, the masks for a single spanning tree across all the brokers encode the spanning tree in the network.) Third, at match time the initialization mask for a given spanning tree (based on the publisher) is refined until the broker can determine whether or not to send a message on each link, that is, until all values in the mask are either Yes or No. These three steps are described in detail in the following three subsections respectively.

A. Annotating the PST

Each broker in the network has a copy of all the subscriptions, organized into a PST as discussed in the previous section, and illustrated in Figure 2. Note that the approach we describe here for computing tree annotations is limited to trees with only equality tests and don't care branches. A more general solution requires the use of a parallel search graph and is not described here to conserve space.Each broker annotates each node of the PST with a trit vector annotation. This annotation vector contains m trits, one per outgoing link from the broker. As mentioned earlier, the trit is Yes when a search reaching that node is guaranteed to match a subscriber reachable through that link, No when a search reaching that node will have no subsearch leading to a subscriber reachable through that

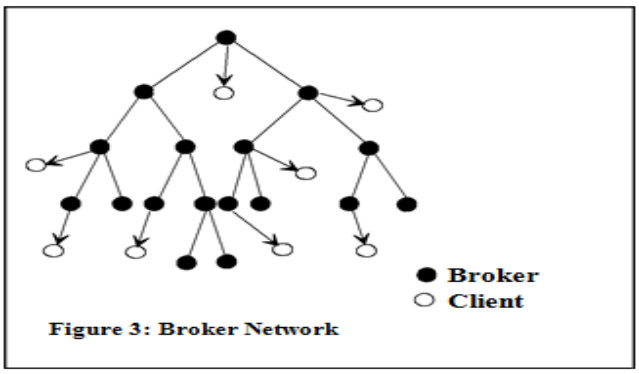


Figure 3: Broker Network

link, and Maybe otherwise.Annotation is a recursive process starting with the leaves of the parallel search tree, which represent the subscriptions. We label each leaf node trit in link position l with Y if one of that leaf node's subscribers is located at a destination reached through link l, and N otherwise. After all the leaves have been annotated, we propagate the annotations back toward the root of the PST using two operators: Alternative Combine and Parallel Combine.Alternative combine is used to combine the annotations from non-* child nodes; Parallel Combine is used to merge the results of the alternative combine operations with the annotation of a child reached by a *-branch.The operators are shown in Figure 4. Intuitively, Alternative Combine takes the least specific result of two annotations. That is, Maybes dominate Yes or No results. Parallel Combine takes the most liberal result of two annotations. That is, Yes dominates Maybe; Maybe dominates No.To compute a node's annotation, Alternative Combine is applied to all children of the node including the one reached through a *-branch. If no *-branch exists, one is included to represent values for which no value branch exists, and an annotation of all No values is added. Parallel Combine is then applied to this result and the *-branch.

An example is shown in Figure 5.

B. Computing The Initialization Mask

We assume that each broker knows the topology of the broker network as well as the best paths between each broker and each destination. To simplify the discussion, we ignore alternative routes for load balancing or recovery from failure and congestion. Instead, we assume that events always follow the shortest path. From this topology information, each broker constructs a routing table

Alternative	Yes	Maybe	No
Yes	Y	M	M
Maybe	M	M	M
No	M	M	N

Parallel	Yes	Maybe	No
Yes	Y	Y	Y
Maybe	Y	M	M
No	Y	M	N

Figure 4: Alternative Combine and Parallel Combine

mapping each possible destination to the link which is the next hop along the best path to the destination.

We also assume that the broker knows the set of spanning trees, only one of which will ever be used by each publisher. In the case where the broker network is acyclic (Figure 3), computation of the spanning tree is straightforward. If the broker topology is not a tree, then computing the spanning tree is more complex. However, even in this case, there will be a relatively small set of different spanning trees. At worst, there will be one spanning tree for each broker that has publisher neighbors and in most practical cases, where the broker network is "tree-like", there will be significantly fewer spanning trees.

Using these best paths and spanning trees, each broker computes the *downstream* destinations for each spanning tree. A destination is downstream from a broker when it is a descendant of the broker on the spanning tree. Based upon the above analysis, each broker associates each unique spanning tree with an *initialization mask*, one trit per link. The trit at link *l* has the value Maybe if at least one of the destinations routable via *l* is a descendant of the broker in the spanning tree; and No if none of the destinations routable via *l* are descendants of the broker[1]. The significance of the mask is that an event arriving at a broker should only be propagated along those links leading to descendant destinations -- that is, those links whose mask bit is M and will eventually be refined to a Y via matching, described below.

C. Matching Events

When an event originating at a publisher is received at a broker, the following steps are taken using the annotated search tree:

i. A mask is created and initialized to the initialization mask associated with the publisher's spanning tree.

ii. Starting with the root node, the mask is *refined* using the trit vector annotation at the current node. During refinement, any M in the mask is replaced by the corresponding trit vector annotation. If the mask is now fully refined --- that is, it has no M trits --- then the search terminates, returning the refined mask. Otherwise, step 3 is executed.

iii. The designated test is performed and, 0, 1, or 2 children are found for continuing the search as mentioned in Section 2. A subsearch is executed at each such child using a copy of the current mask.

On the return of each subsearch, all Maybe trits in the current mask for which a Yes trit exists in the subsearch mask, are converted to Yes trits. After all the children have been searched, the remaining Maybe trits in the current mask are made No trits. The current mask is returned.

iv. The top-level search terminates and sends a copy of the event to all links corresponding to Yes trits in the returned mask.

This concludes the description of the link matching algorithm.

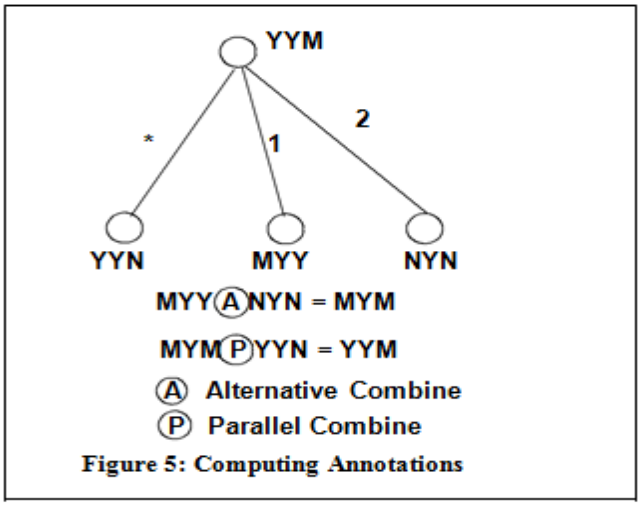


Figure 5: Computing Annotations

IV. Implementation and Performance

We have implemented the matching algorithms described above and tested them on a simulated network topology as well as on a real LAN, as explained in the following two subsections respectively.

A. Simulation Results

The goals of our simulations were twofold

i. To measure the network loading characteristics of the link matching protocol and compare it to that of the flooding protocol.

ii. To measure the processing time taken by the link matching algorithm at individual broker nodes and compare it to that of centralized matching (i.e., the non-trit matching algorithm described in Section 2).

i. Simulation Setup

The simulated broker network topology is shown in Figure 6. The topology has 39 brokers and 10 subscribing clients per broker, each client with potentially multiple subscriptions. In addition, there is an unspecified number of publishing clients -- three of these publishers, shown as P1, P2, and P3 in the figure, publish events that are tracked by the simulator and the rest simply load the brokers by publishing messages that take up CPU time at the brokers.

As shown in Figure 6, the 39 brokers form three trees of 13 brokers each. The root of each of these three trees are connected to the roots of the other two. Also, as shown, there are a small number of lateral links between non-root nodes in the trees to allow messages from some publishers to follow a different path than other publishers. This topology is intended to model a real-world wide-area network with each of the three rooted trees distributed far

[1] *In some cases, where some destinations reachable through a link downstream on some spanning trees and are not on others, the search may be optimized by splitting the link into two or more "virtual" links.*

from each other (intercontinental), but the brokers within a tree closer to each other (interstate). The top-level brokers are modeled to have a one-way hop delay of about 65 ms, links from them to their next level neighbors is 25ms, the third level hop delay is about 10ms, and the hop delay to clients is 1ms.

The broker network simulates an information space with several control parameters, such as the number of attributes in the event schema, the number of values per attribute and the number of factoring levels (i.e., the preferred attributes of Section 2.1). Subscriptions are generated randomly, but one of the control parameters is the probability that each attribute is a * (i.e., don't care). For non-* attributes, the values are generated according to a zipf distribution. In addition, we simulate "locality of interest" in subscriptions by having subscribers within each subtree of the broker topology have similar distributions of interested values whereas subscriptions across from the other two subtrees have different distributions.

Events are also generated randomly, with attribute values in a zipf distribution. Events arrive at the publishing brokers according to a Poisson distribution. The mean arrival rate of published events, which is a key parameter, is controlled by a user specified parameter.

In the simulation, time is measured in "ticks" of a virtual clock, with each tick corresponding to about 12 microseconds. The virtual clock, used only for simulation purposes, is implemented as synchronized brokers' clocks. Each event carries with it its "current" virtual time from the beginning of the simulation. An event spends time traversing a link ("hop delay"), waiting at an incoming broker queue, getting matched, and being sent (software latency of the communication stack).

ii. Network Loading Results

As mentioned earlier, the purpose of this simulation run was to determine, for the link matching and the flooding protocols, the event publish rate at which the broker network becomes "overloaded" (or congested), for a varying number of subscriptions. A broker is overloaded when its input message queue is growing at a rate higher than the broker processor can handle.

This simulation run was performed with the following parameters. The event schema has 10 attributes (with 2 attributes used for factoring), and each attribute has 5 values. The subscriptions are generated randomly in such a way that the first attribute is non-* with probability 0.98, and this probability decreases at the rate of 85% as we go from the first to the last attribute. This means that subscriptions are very selective -- on average, each event matches only about 0.1% of subscriptions. The number of events published is 500.

The results from the simulation run are shown in Chart 1. The chart shows that a broker network running the flooding protocol saturates at significantly lower event publish rates than the link matching protocol for any number of subscriptions. In particular, when each event is destined to only a small percentage of all clients, link matching dramatically outperforms flooding. In the case where events are distributed quite widely, the difference is not as great, since most links are used to distribute events in the link matching protocol. This result illustrates that link matching is well-suited to the type of selective multicast that is typical of pub/sub systems deployed on a WAN.

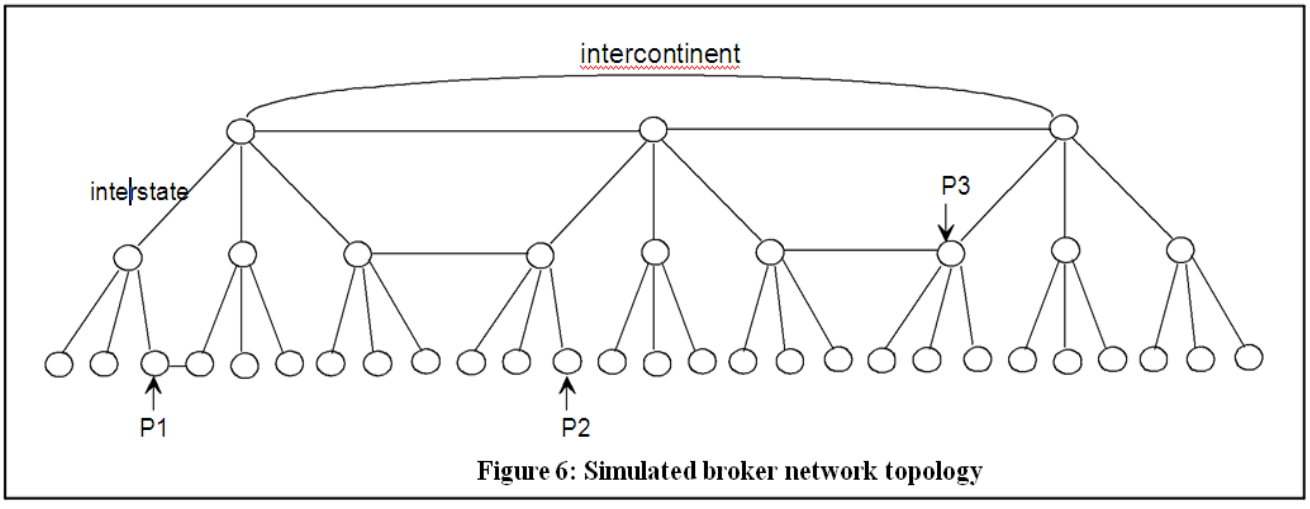
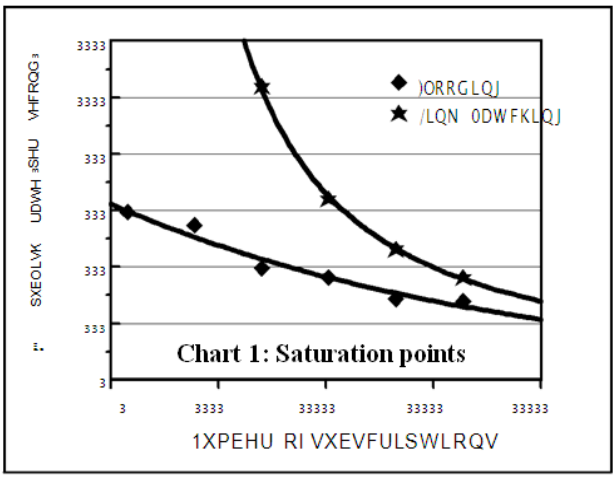
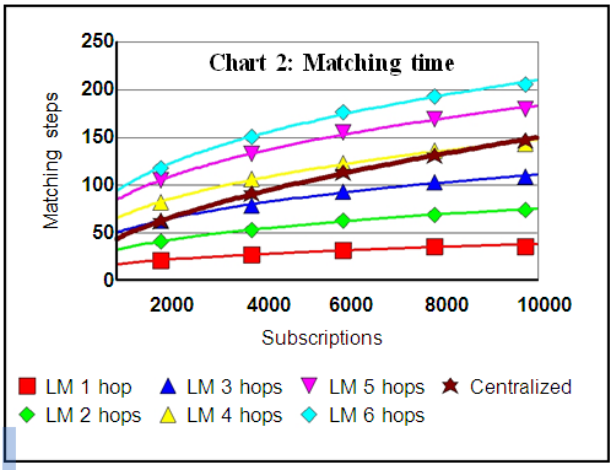


Figure 6: Simulated broker network topology





iii. Matching Time Results

As mentioned earlier, the purpose of this simulation run was to measure the cumulative processing time taken by the link matching algorithm and the centralized (non-trit) matching algorithm. The processing time taken per event in the link matching algorithm is the sum of the times for all the partial matches at intermediate brokers along the way from publisher to subscriber.

This simulation run was performed with the following parameters. The event schema has 10 attributes (with 3 attributes used for factoring), and each attribute has 3 values. The subscriptions are generated randomly in such a way that the first attribute is non-* with probability 0.98, and this probability decreases at the rate of 82% as we go from the first to the last attribute. Again, this means that subscriptions are very selective -- on average, each event matches only about 1.3% of subscriptions. The number of events published is 1000.

The results from the simulation run are shown in Chart 2. For the link matching algorithm, six lines, —LM1 hop" through —LM6 hops", are shown -- these correspond to the number of hops an event had to traverse on its way from a publishing broker to a subscriber. On the Y axis, the chart shows the number of —matching steps" performed on average. A matching step is the visitation of a single node in the matching tree. Although our current implementation has traded off time efficiency in favor of space efficiency, we estimate that a time efficient implementation can execute a matching step in the order of a few microseconds.

The chart shows that the cumulative matching steps for up to four hops using the link matching algorithm is not more than the number of matching steps taken by the centralized algorithm. For more than four hops the link matching algorithm takes more matching steps, however the link matching protocol is still a better choice over the centralized algorithm because (1) the extra processing time for link matching (of the order of much less than 1ms) is insignificant compared to network latency (of the order of tens of ms), (2) the gain in latency to regional publishers and subscribers obtained by distributing brokers is significant, and (3) for really large numbers of subscribers (i.e., much beyond 10000), the slopes of the lines in Chart 2 indicate that centralized matching may take more steps than link matching.

B. System Prototype

We have implemented the matching algorithms in a network of broker nodes where brokers are connected using a specified topology. A broker network may implement multiple information spaces by specifying an event schema (one per information space) defining the type of information contained in each event. Clients subscribe to an information space by first connecting to a broker node, then providing subscription information which includes a predicate expression of event attributes. This section describes the implementation of such a broker node.

As illustrated in Figure 7, each broker node consists of a matching engine, client and broker protocols, a connection manager and a transport layer. The matching engine which implements one of the matching algorithms described earlier, consists of a subscription manager, and an event parser. A subscription manager receives a subscription from a client, parses the subscription expression, and adds the subscription to the matching tree. An event parser first parses a received event, then un-marshals it according to the pre-defined event schema. The matchine engine then uses the implemented matching algorithm to get a list of subscibers interested in the un-marshaled event.

The broker to client protocol is implemented by the client protocol object, whereas the broker to broker protocol is implemented using the broker protocol object. These protocol objects are robust enough to handle transient failures of connections by maintaining an event log per client. Once a client re-connects after a failure, the client protocol object delivers the events received while the client was dis-connected. A garbage collector periodically cleans up the log. The connection manager object maintains the connections to clients and the other brokers in the network.

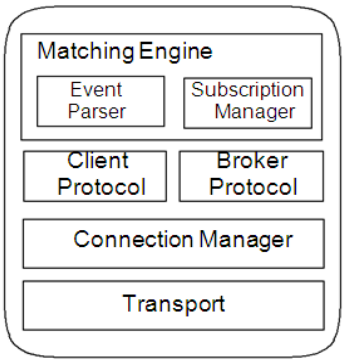


Figure 7: Broker Components

The transport layer sends and receives messages to and from clients and other brokers in the network. To improve scalability, it implements an asynchronous "send" operation by maintaining a set of outgoing queues, one per connection. A broker thread sends a message by en-queueing it in the appropriate queue. A pool of sending threads is responsible for monitoring these queueues for outgoing messages, and sending them to destinations using the underlying network protocol.

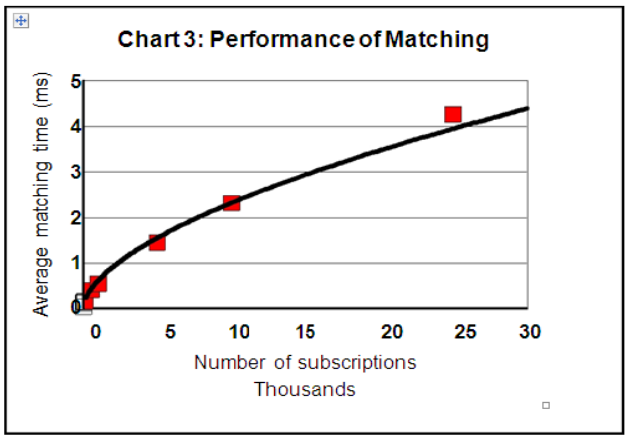
Currently, broker nodes are implemented in Java using TCP/IP as the network protocol. In an experimental setup where a 200 MHz pentium pro PC is used as a broker node, and low end PCs (using 133 MHz pentium processors) are used as clients connected using a 16MB token ring network, the current implementation of the broker can deliver upto 14,000 events/sec. Also, as shown in Chart 3 for the pure matching algorithm, brokers can perform matching very quickly, at the rate of about 4ms for 25,000 subscribers. In fact, our matching algorithms are so efficient that the transport system and network costs of a broker outweigh the cost of matching at a broker.

V. Related Work

As mentioned earlier, alternatives to the link matching approach were either to (1) first compute a destination list for events by matching at or near the publisher and then distributing the event using the distribution list, or (2) to multicast the event to all subscribers which would then filter the event themselves.

Computing a destination list is a good approach for small systems involving only a few subscribers. For these cases, the matching algorithm presented in section 2 provides a good solution. However, scalability is essential if content-based systems are to fill the same infrastructure requirements as subject-based publish/subscribe systems. In cases where destination lists may grow to include hundreds or thousands of destinations, the match-first approach becomes impractical.

Multicasting an event and then filtering also has its disadvantages. Lack of scalability and an inability to exploit locality was shown for the flooding approach for



event distribution. Flooding is a good approximation of the broadcast approach since most WAN multicast techniques require the use of a series of routers or bridges connecting LAN links. IP multicast [5][1] allows subscriptions to a subrange of possible IP addresses known as class D addresses. Subscriptions to these groups is propagated back through the network routers implementing IP. Pragmatic General Multicast [16] has been proposed as an internet-wide multicast protocol with a higher level of service. This protocol is an extension of IP multicast that provides "TCP-like" reliability, and therefore is also reliant on multicast-enabled routers. A mechanism for multicast in a network of bridge-connected LANs is proposed in [7]. In this approach, members of a group periodically broadcast to an all-bridge group their membership in a multicast group. Bridges note these messages and update entries in a multicast table, including an expiration time.

The content-based subscription systems that have been developed do not yet address wide-area, scaleable event distribution, i.e. although they are content-based subscription systems, they are not content-based routing systems. SIENA allows content-based subscriptions to a distributed network of event servers (brokers) [6]. SIENA filters events before forwarding them on to servers or clients. However, a scaleable matching algorithm for use at each server has not been developed. The Elvin system [14] uses an approach similar to that used in SIENA. Publishers are informed of subscriptions so that they may "quench" events (not generate events) for which there are no subscribers. In [14], plans are discussed for optimizing Elvin event matching by integrating an algorithm similar to the parallel search tree. This algorithm, presented in [8], converts subscriptions into a deterministic finite automata for matching. However, no plans for optimizations for broker links (such as our optimization through trit annotation) are discussed.

Another algorithm for optimizing matching is discussed in [9]. At analysis time, one of the tests aij of each subscription is chosen as the gating test; the remaining tests of the subscription (if any) are residual tests. At matching time, each of the attributes aj in the event being matched is examined. The event value vj is used to select those subscriptions i whose gating tests include aij = vj. The residual tests of each selected subscription are then

evaluated: if any residual test fails, the subscription is not matched; if all residual tests succeed, the subscription is matched. Our parallel search tree performs this type of test for each attribute, not just a single gating test attribute.

One outlet for the work presented in this paper could be through Active Networks [17]. Active Networks have been touted as a mechanism for eliminating the strong dependence of route architectures on Internet standards. Active Networks allow the dynamic inclusion of code either at routers or by replacing passive packets with active code. The SwitchWare project [3] follows the former approach and is most appropriate to the type of router customization proposed in this paper. With SwitchWare, digitally signed type-checked modules may be loaded into network routers. Our matching and multicasting component could be one such module.

VI. Conclusions

In this paper, we have presented a new multicast technique for content-based publish/subscribe systems known as link matching. Although several publish/subscribe systems have begun to support content-based subscription, the novel contribution of link matching is that *routing* is based on a hop-by-hop partial matching of published events. The link matching approach allows distribution of events to a large number of information consumers distributed across a WAN without placing an undo load on the network. The approach also exploits locality of subscriptions.

We evaluate how an implementation of content-based routing protocol performs by showing that a broker network stays up while running the link matching algorithm whereas brokers get overloaded for the same event arrival rate running the flooding algorithm, since brokers have larger numbers of events to process in the flooding case. We also describe a broker implementation that can handle message loads of up to 14000 events per second on a 200 MHz Pentium PC. This shows that content-based routing using link matching supports a more general and flexible form of publish-subscribe while admitting a highly efficient implementation.

Future work is concentrating on further validation of our approach to content-based routing. We are currently working to deploy our content-based routing brokers on a large private network. This will allow us to conduct system tests under actual application loads. Sample applications will include some from the financial trading and process control domains. In addition to these system tests, we are also continuing work with our simulator to examine different types of messaging loads. In particular, since many publish/subscribe applications exhibit peak activity periods, we are examining how our protocol performs with bursty message loads.

VII. Reference

1) Lorenzo Aguilar. —Datagram Routing for Internet Multicasting," ACM Computer Communications Review, 14(2), 1984. pp. 48-63.
2) Marcos Aguilera, Rob Strom, Daniel Sturman, Mark Astley, Tushar Chandra. 1998. Matching Events in a Content-Based Subscription System. Upcoming IBM Technical Report, available from http://www.research.ibm.com/gryphon.
3) Scott Alexander et al., —The SwitchWare Active Network Architecture," IEEE Network Special Issue on Active and Controllable Networks, July 1998, Vol. 12, No. 3. pp. 29--36.
4) K. P. Birman. —The process group approach to reliable distributed computing," pages 36-53, Communications of the ACM, Vol. 36, No. 12, Dec. 1993.
5) Uyless Black. TCP/IP & Related Protocols, Second Edition. McGraw-Hill, 1995. pp. 122-126.
6) Antonio Carzaniga, David S. Rosenblum, and Alexander L. Wolf. —Design of a Scalable Event Notification Service: Interface and Architecture," unpublished. Available from http://www.cs.colorado.edu/users/carzaniga/siena/index.html
7) Stephen E. Deering. —Multicast Routing in InterNetworks and Extended LANs," ACM Computer Communications Review, 18(4), 1988 . pp. 55-64.
8) John Gough and Glenn Smith. —Efficient Recognition of Events in a Distributed System," Proceedings of ACSC-18, Adelaide, Australia, 1995.
9) Eric N. Hanson, Moez Chaabouni, Chang-Ho Kim, Yu-Wang Wang. —A predicate Matching Algorithm for Database Rule Systems," pages 271-280, SIGMOD 1990, Atlantic City N. J., May 23-25 1990.
10) Shivakant Mishra, Larry L. Peterson, and Richard D. Schlichting. Consul: A Communication Substrate for Fault-Tolerant Distributed Programs, Dept. of computer science, The University of Arizona, TR 91-32, Nov. 1991.
11) Object Management Group. CORBA services: Common Object Service Specification. Technical report, Object Management Group, July 1998.
12) Brian Oki, Manfred Pfluegl, Alex Siegel, Dale Skeen. —The Information Bus - An Architecture for Extensible Distributed Systems," pages 58-68, Operating Systems Review, Vol. 27, No. 5, Dec. 1993.
13) David Powell (Guest editor). —Group Communication", pages 50-97, Communications of the ACM, Vol. 39, No. 4, April 1996.
14) Bill Segall and David Arnold. —Elvin has left the building: A publish/subscribe notification service with quenching," Proceedings of AUUG97, Brisbane, Austrailia, September, 1997.
15) Dale Skeen. Vitria's Publish-Subscribe Architecture: Publish-Subscribe Overview, http://www.vitria.com/

16) Tony Speakman, Dino Farinacci, Steven Lin, and Alex Tweedly. ―PGM Reliable Transport Protocol,‖ IETF Internet Draft. August 24, 1998.
17) Tennenhouse, J. Smith, W. D. Sincoskie, D. Wetherall, G. Minden. ―A Survey of Active Network Research,‖ IEEE Communications Magazine. January, 1997, Vol. 35, No. 1. pp. 80--86.

Implementation Of A Radial Basis Function Using VHDL

GJCST Classification
I.2.6. K.3.2

[1]D.C. Dhubkarya , [2]Deepak Nagariya , [3]Richa Kapoor

***Abstract*- This paper presents the work regarding the implementation of neural network using radial basis function algorithm on very high speed integrated circuit hardware description language (VHDL). It is a digital implementation of neural network. Neural Network hardware has undergone rapid development during the last decade. Unlike the conventional von-Neumann architecture that is sequential in nature, Artificial Neural Networks (ANNs) Profit from massively parallel processing. A large variety of hardware has been designed to exploit the inherent parallelism of the neural network models.**

The radial basis function (RBF) network is a two-layer network whose output units form a linear combination of the basis function computed by the hidden unit & hidden unit function is a Gaussian. The radial basis function has a maximum of 1 when its input is 0. As the distance between weight vector and input decreases, the output increases. Thus, a radial basis neuron acts as a detector that produces 1 whenever the input is identical to its weight vector.

Keywords- RBF, training algorithm, weight, block RAM, FPGA.

I. INTRODUCTION

Artificial Neural Networks (ANNs) are non-linear mapping structures based on the function of the human brain. They are powerful tools for modeling, especially when the underlying data relationship is unknown. ANNs can identify and learn correlated patterns between input data sets and corresponding target values. ANNs imitate the learning process of the human brain and can process problems involving non-linear and complex data even if the data are Imprecise and noisy. An ANN is a computational structure that is inspired by observed process in natural networks of biological neurons in the brain. It consists of simple computational units called neurons, which are highly interconnected. ANNs have become the focus of much attention, largely because of their wide range of applicability and the ease with which they can treat complicated problems. ANNs are parallel computational models comprised of densely interconnected adaptive processing units.[1] These networks are fine-grained parallel Implementations of nonlinear static or dynamic systems. A very important feature of these networks is their adaptive nature, where ―learning by example" replaces ―programming" in solving problems. This feature makes such computational models very appealing in application domains where one has little or incomplete understanding of the problem to be solved but where training data is readily Available. ANNs are now being increasingly recognized in the area of classification and prediction, where regression model and other related statistical techniques have traditionally been employed.

[1]*D.C. Dhubkarya , UPTU, Lucknow , BIET Jhani, dcd3580@yahoo.com*
[2]*Deepak Nagariya , UPTU, Lucknow , BIET Jhansi , deepaknagaria@gmail.com*
[3]*Richa Kapoor, UPTU,Lucknow , SIT,Mathura, richakapoor11@gmail.com*

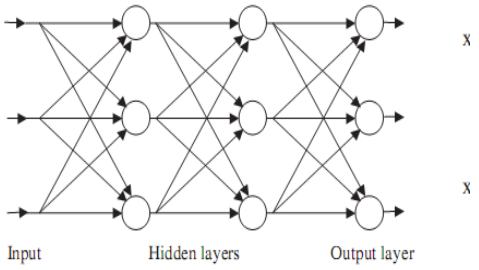


Fig. 1: Schematic representation of neural network

II. OVERVIEW OF RADIAL BASIS FUNCTION (RBF) NETWORKS"

Radial basis function (RBF) neural network consist of three layers, an input, a hidden and an output. It has a feed forward structure consisting of a single hidden layer of J locally tuned units, which are fully interconnected to an output layer of L linear units. All hidden units simultaneously receive the n-dimensional real-valued input vector X .The hidden-unit outputs are not calculated using the weighted-sum mechanism/sigmoid activation; rather each hidden-unit output $\varphi_i(X)$ is obtained by closeness of the input X to an n-dimensional parameter vector Cj associated with the jth hidden unit. [13] The response characteristics (activation function) of the jth hidden unit (j = 1, 2, .. J) is assumed as,

$$\varphi_i(X) = \exp\left\{-\frac{\|X - C_i\|^2}{2\sigma_i^2}\right\}.$$

The Parameter σj is the width of the receptive field in the input space from unit j. This implies that $\varphi_i(X)$ has an appreciable value only when the distance $\|X - C_i\|$ is smaller than the width σj..

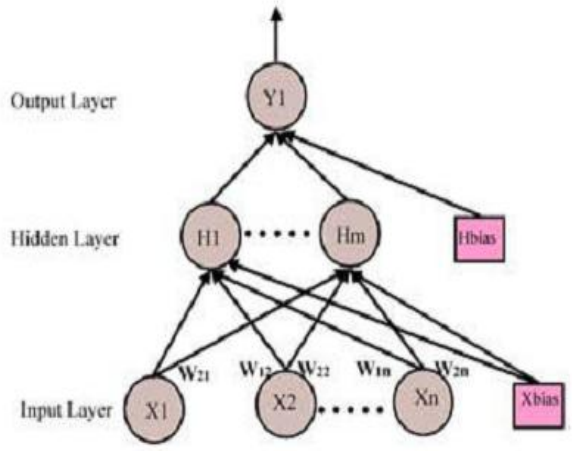


Fig2: Feed Forward Neural Network

RBF networks are best suited for approximating continuous or piecewise continuous real-valued mapping. $f: R^n \rightarrow R^L$, where n is sufficiently small. These approximation problems include classification problems as a special case. In the present work we have used a Gaussian basis function for the hidden units. RBF networks have been successfully applied to a large diversity of applications including interpolation, chaotic time-series modeling, system identification, control engineering, electronic device parameter modeling, channel equalization, speech recognition, image restoration, shape- from-shading, 3-D object modeling, motion estimation and moving object segmentation etc [7].

III. TRAINING OF RBF NEURAL NETWORKS

By means of training, the neural network models the underlying function of a certain mapping. In order to model such a mapping we have to find the network weights and topology. There are two categories of training algorithms: supervised and unsupervised. In supervised learning, the model defines the effect one set of observations, called inputs, has on another set of observations, called outputs. In other words, the inputs are assumed to be at the beginning and outputs at the end of the causal chain. The models can include mediating variables between the inputs and outputs. In unsupervised learning, all the observations are assumed to be caused by latent variables, that is, the observations are assumed to be at the end of the causal chain. In practice, models for supervised learning often leave the probability for inputs undefined.[1]

RBF networks are used mainly in supervised applications. In a supervised application, we are provided with a set of data samples called training set for which the corresponding network outputs are known.

RBF networks are trained by

i. deciding on how many hidden units there should be
ii. deciding on their centres and the sharpnesses (standard deviation) of their Gaussians
iii. Training up the output layer.

In training phase, a set of training instances is given. A feature vector typically describes each training instances. It is further associated with the desired outcome, which is further represented by a feature vector called output vector. Starting with some random weight setting, the neural net is trained to adapt itself by changing weight inside the network according to some learning algorithm. When the training phase is complete the weights are fixed. The network propagates the information from the input towards the output layer. When propagation stops, the output units carry the result of the inferences.

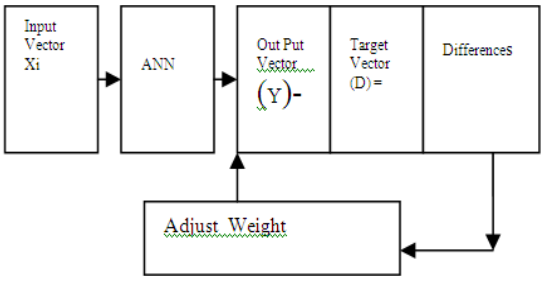


Fig 3: Block Diagram

Learning law describes the weight vector for the ith processing unit at time instant (t+1) in terms of the weight vector at time instant (t) as follows;

$$w_i(t+1) = w_i(t) +/- \Delta w_i(t),$$

Where $\Delta w_i(t)$ is the change in the weight vector. The Networks adapt change the weight by an amount proportional to the difference between the desired output and the actual output. As an equation:

$\Delta W_i = \eta * (D-Y).Xi$

Here E=D-Y

The perceptron learning rule can be written more succinctly in terms of the error E and the change to be made to the weight vector ΔW_i

CASE 1- If E = 0, then make a change ΔW equal to 0.

CASE 2- If E = +, then make a change

$$w_i(t+1) = w_i(t) + \Delta w_i(t)$$

CASE 3-. If E = -1, then make a change

$$w_i(t+1) = w_i(t) - \Delta w_i(t),$$

Where η is the learning rate, D is the desired output, Y is the actual output, and Ii is the ith input. The weights in an ANN, similar to coefficients in a regression model, are adjusted to solve the problem presented to ANN. Learning or training is term used to describe process of finding values of these weights. Two types of learning with ANN are supervised and unsupervised learning. An important issue concerning supervised learning is the problem of error convergence, i.e. the minimization of error between the desired and computed unit values. The aim is to determine a set of weights which minimizes the error.

IV. IMPLEMENTATION

Parameter

η=0.8;

Weight vector=8;

Input Vector=8;
Target value=1;

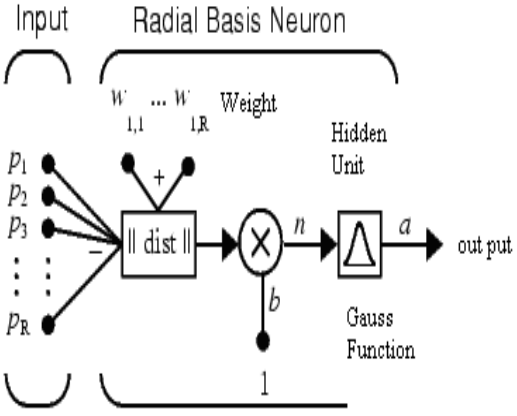


Fig: 4 Neuron Model

$f(\bar{x}) = e^{-Cu^2}$ where C is a constant and $u = \sqrt{\sum_i ((x_i - \omega_i)^2)}$

Since in the transfer function the u is squared, the square-root in u is unnecessary (especially in the hardware), and the function becomes

$$f(\bar{x}) = e^{-Cv} \quad , \quad v = \sum_i ((x_i - \omega_i)^2)$$

(The `|| dist ||` box in this figure accepts the input vector **p** and the single row input weight matrix.

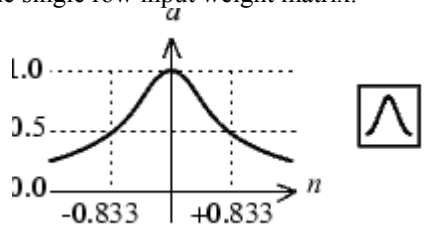


Fig 5: Radial Basis Function

we can understand how this network behaves by following an input vector **p** through the network to the output **a**. we present an input vector to such a network, each neuron in the radial basis layer will output a value according to how close the input vector is to each neuron's weight vector. Thus, radial basis neurons with weight vectors quite different from the input vector **p** have outputs near zero. These small outputs have only a negligible effect on the linear output neurons.

In contrast, a radial basis neuron with a weight vector close to the input vector **p** produces a value near 1. If a neuron has an output of 1, its output weights in the second layer pass their values to the linear neurons in the second layer.

In fact, if only one radial basis neuron had an output of 1, and all others had outputs of 0's (or very close to 0), the output of the linear layer would be the active neuron's output weights. This would, however, be an extreme case. Typically several neurons are always firing, to varying degrees. if it's output is not 1 then weight is adjusted according to training algorithm used & weight is updated till desired valued matched with target value.

V. Simulation Results

The proposed design was coded in VHDL. It was functionally verified by writing a test bench and simulating it using ISE simulator and synthesizing it on Spartan 3A using Xilinx ISE 9.2i.

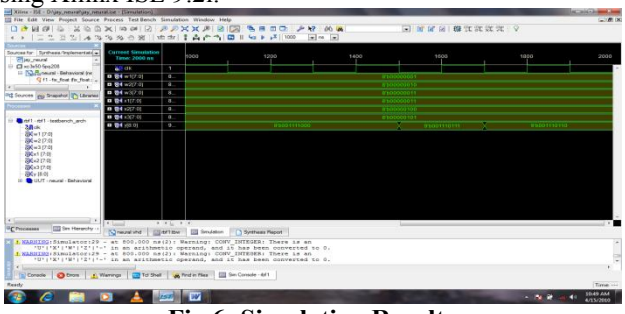


Fig 6: Simulation Results

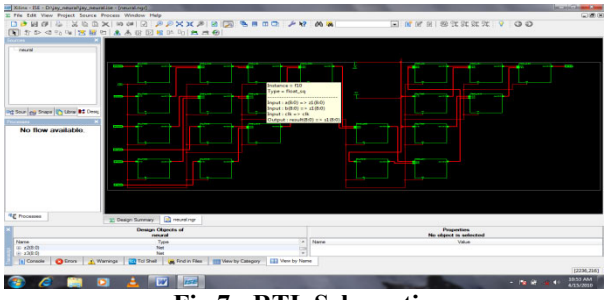


Fig 7 : RTL Schematic

HDL Synthesis Report

Output File Name —neural"
Output Format: NGC
Target Device: xc3s50-5-pq208
Number of Slices: 1000 out of 768
Number of Slice F/F: 211 out of 1536
Number of 4 input LUTs: 1713 out of 1536
Number of IOs: 58
Number of bonded IOBs: 58 out of 124
Number of GCLKs: 1 out of 8 12%

VI. References

1) Adrian G. Bors —Introduction of the Radial Basis Function (RBF) Networks" Department of Computer Science University of York YO10 5DD, UK.
2) Broomhead, D. S., Jones, R., McWhirter, J. G., Shepherd, T. J., (1990) —Systolic array for nonlinear multidimensional interpolation using radial basis functions," Electronics Letters, vol. 26, no. 1, pp. 7-
3) Bors, A.G., Pitas, I., (1996) —Median radial basis functions neural network," IEEE Trans. on Neural Networks, vol. 7, no. 6, pp. 1351-1364.
4) Casdagli, M. (1989) Nonlinear prediction of chaotic time series," Physica D, vol. 35, pp. 335-356.

5) Chen, S., Cowan, C. F. N., Grant, P. M. (1991) —Orthogonal least squares learning Algorithm for radial basis function networks," IEEE Trans. On Neural Networks, vol. no. 2, pp. 302-309.
6) Douglas L Perry (2006), —VHDL Programming by Example", McGraw- Hill.
7) F. Belloir, A. Fache and A. Billat (1998), —A New Construction Algorithm of efficient Radial Basis Function Neural Net Classifier and its Application to codes Identification",Ardenne, B.P. 1039, 51687.
8) Igelnik, B., Y.-H. Pao, (1995) —Stochastic choice of radial basis functions in adaptive function approximation and the functional-link net," IEEE Trans. on Neural Networks, vol. 6, no. 6, pp. 1320-1329.
9) Karayiannis,N.B.(1999) Reformulated radial basis neural networks trained by gradient descent," IEEE Trans. on Neural Networks, vol. 10, no. 3, pp. 657-671.
10) Kohonen,T.K.,(1989)Self-organization and associative memory. Berlin: Springer- Verlag.
11) Karen Parnell & Nick Mehta (2002), —Programmable Logic Design Quick Start Hand Book", Xilinx.
12) Musavi, M.T., Ahmed, W., Chan, K.H., Faris,K.B., Hummels, D.M., (1992) —On the training of radial basis function classifiers," Neural Networks, vol. 5, pp. 595-603.
13) P. Venkatesan* and S. Anitha,—Application of a radial basis function neural network for diagnosis of diabetes mellitus" Tuberculosis Research Centre, ICMR, Chennai 600 031, India
14) Satish Kumar —Neural Networks —A classroom approach, TMH.
15) Yuehui Chen1, Lizhi Peng1, and Ajith Abraham (2004)," Hierarchical Radial Basis Function NeuralNetworks for Classification Problems" International Journal of Neural Systems, 14(2):125-137.

Analysis And Implementation Of Data Mining Techniques , Using Naive-Bayes Classifier And Neural Networks

GJCST Classification
H.2.8 I.2.6

[1]Sudheep Elayidom.M, [2]Sumam Mary Idikkula, [3]Joseph Alexander

***Abstract*--Taking wise career decision is so crucial for anybody for sure. In modern days there are excellent decision support tools like data mining tools for the people to make right decisions. This paper is an attempt to help the prospective students to make wise career decisions using technologies like data mining. In India technical manpower analysis is carried out by an organization named NTMIS (National Technical Manpower Information System), established in 1983-84 by India's Ministry of Education & Culture. The NTMIS comprises of a lead centre in the IAMR, New Delhi, and 21 nodal centres, located at different parts of the country. The Kerala State Nodal Centre is located in the Cochin University of Science and Technology. Last 4 years information is obtained from the NODAL Centre of Kerala State (located in CUSAT, Kochi, India), which stores records of all students passing out from various technical colleges in Kerala State, by sending postal questionnaire. Analysis is done based on Entrance Rank, Branch, Gender (M/F), Sector (rural/urban) and Reservation (OBC/SC/ST/GEN). Using this data, data mining models like Naïve Bayes classifier and neural networks are built, tested and used to predict placement chances, given inputs like rank, sector, category and sex.**

Keywords- Data mining, WEKA, Neural networks, Naïve Bayes classifier, Confusion matrix.

I. INTRODUCTION

The popularity of subjects in science and engineering in colleges around the world is up to a large extent dependent on the viability of securing a job in the corresponding field of study. Appropriation of funding of students from various sections of society is a major decision making hurdle particularly in the developing countries.

An educational institution contains a large number of student records. This data is a wealth of information, but is too large for any one person to understand in its entirety. Finding patterns and characteristics in this data is an essential task in education research.This type of data is presented to decision makers in the State Government in the form of tables or charts, and without any substantive analysis, most analysis of the data is done according to individual intuition, or is interpreted based on prior research. This paper is an attempt to scientifically analyze the trends of placements keeping in account of details like Entrance Rank, Sex, Category and Reservation using Naive Bayes classifier and Neural Networks.

The data preprocessing for this problem has been described in detail in articles [1] & [9], which are papers published by the same authors. The problem of placement chance prediction may be implemented using decision trees. [4] Surveys a work on decision tree construction, attempting to identify the important issues involved, directions which the work has taken and the current state of the art. Studies have been conducted in similar area such as understanding student data as in [2]. There they apply and evaluate a decision tree algorithm to university records, producing graphs that are useful both for predicting graduation, and finding factors that lead to graduation. It's always been an active debate over which engineering branch is in demand .So this work gives a scientific solution to answer these. Article [3] provides an overview of this emerging field clarifying how data mining and knowledge discovery in databases are related both to each other and to related fields, such as machine learning, statistics, and databases. [5] Suggests methods to classify objects or predict outcomes by selecting from a large number of variables, the most important ones in determining the outcome variable. The method in [6] is used for performance evaluation of the system using confusion matrix which contains information about actual and predicted classifications done by a classification system. [7] & [8] suggest further improvements in obtaining the various measures of evaluation of the classification model.

II. DATA

The data used in this project is the data supplied by National Technical Manpower Information System (NTMIS) via Nodal center. Data is compiled by them from feedback by graduates, post graduates, diploma holders in engineering from various engineering colleges and polytechnics located within the state during the year 2000-2003. This survey of technical manpower information was originally done by the Board of Apprenticeship Training (BOAT) for various individual establishments. A prediction model is prepared from data during the year 2000-2002 and tested with data from the year 2003.

Sudheep Elayidom
Computer Science and Engineering Division, School Of Engineering
Cochin University of Science and Technology, Kochi, India
+91-04842463306,sudheepelayidom@hotmail.com
Sumam Mary Idikkula
Department of Computer Science
Cochin University of Science and Technology, Kochi, India
+91-04842577605, sumam@cusat.ac.in
Joseph Alexander
Project Officer, Nodel Center
Cochin University of Science and Technology, Kochi, India
+91-04842862406, josephalexander@cusat.ac.in

III. Problem Statement

Modeling and predicting the chances of placements in colleges keeping account of details like Rank, Gender, Branch, Category, Reservation and Sector. It also analyses and compares the performances of Naive Bayes classifier and neural networks for this problem.

IV. Concepts Used

A. Data Mining

Data mining is the principle of searching through large amounts of data and picking out interesting patterns. It is usually used by business intelligence organizations, and financial analysts, but it is increasingly used in the sciences to extract information from the enormous data sets generated by modern experimental and observational methods. A typical example for a data mining scenario may be ―In mining analysis if it is observed that people who buy butter tend to buy bread too then for better business results the seller can place butter and bread together."

B. Naive Bayes classifier

In simple terms, a naive Bayes classifier assumes that the presence (or absence) of a particular feature of a class is unrelated to the presence (or absence) of any other feature.
Depending on the precise nature of the probability model, Naive Bayes classifiers can be trained very efficiently in a supervised learning setting. In many practical applications, parameter estimation for Naive Bayes model uses the method of maximum likelihood. Despite its simplicity, Naive Bayes can often outperform more sophisticated classification methods. Recently, careful analysis of the Bayesian classification problem has shown that there are some theoretical reasons for the apparently unreasonable efficacy of naive Bayes classifiers.
An advantage of the naive Bayes classifier is that it requires a small amount of training data to estimate the parameters (means and variances of the variables) necessary for classification. Because independent variables are assumed, only the variances of the variables for each class need to be determined and not the entire covariance matrix.
The classifier is based on Bayes theorem, which is stated as

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Each term in Bayes' theorem has a conventional name:
*P (A) is the prior probability or marginal probability of A. It is "prior" in the sense that it does not take into account any information about B.
*P (A|B) is the conditional probability of A, given B. It is also called the posterior probability because it is derived from or depends upon the specified value of B.
*P (B|A) is the conditional probability of B given A.
*P (B) is the prior or marginal probability of B, and acts as a normalizing constant.
Bayes' theorem in this form gives a mathematical representation of how the conditional probability of event A given B is related to the converse conditional probability of B given A.

C. Weka

WEKA is a collection of machine learning algorithms for data mining tasks. WEKA contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization. It is also well-suited for developing new machine learning schemes.
The main strengths of WEKA are that it is
- Very portable because it is fully implemented in the Java programming language and thus runs on almost any modern computing platform,
-contains a comprehensive collection of data pre-processing and modelling techniques, and is easy to use by a novice due to the graphical user interfaces it contains.
WEKA's main user interface is the Explorer, but essentially the same functionality can be accessed through the component-based Knowledge Flow interface and from the command line. There is also the Experimenter, which allows the systematic comparison of the predictive performance of WEKA's machine learning algorithms on a collection of datasets.

D. Confusion Matrix

A confusion matrix is a visualization tool typically used in supervised learning (in unsupervised learning it is typically called a matching matrix). It is used to represent the test result of a prediction model. Each column of the matrix represents the instances in a predicted class, while each row represents the instances in an actual class. One benefit of a confusion matrix is that it is easy to see if the system is confusing two classes (i.e. commonly mislabelling one as another).When a data set is unbalanced (when the number of samples in different classes vary greatly) the error rate of a classifier is not representative of the true performance of the classifier. This can easily be understood by an example: If there are for example 990 samples from class A and only 10 samples from class B, the classifier can easily be biased towards class A. If the classifier classifies all the samples as class A, the accuracy will be 199%. This is not a good indication of the classifier's true performance. The classifier has a 100% recognition rate for class A but a 0% recognition rate for class B.

E. Data Pre-Processing

The Initial database provided by Nodal Center was loaded to MS Excel and converted to CSV files (Comma Separated files) .This file was loaded in WEKA Knowledge Flow interface and converted into ARFF files (Attribute-Relation File Format). The individual database files(DBF format) for the years 2000-2003 were obtained and one containing records of students from the year 2000-2002 and another for year 2003, were created.
List of attributes extracted:

RANK: Rank secured by candidate in the engineering entrance exam. Range: 1-25
C*ATEGORY*: Social background. Range: {General, Scheduled Cast, Scheduled Tribe, Other Backward Class}
SEX : Range{Male, Female}
SECTOR : Range {Urban, Rural}
BRANCH : Range{A-J}
PLACEMENT: Indicator of whether the candidate is placed.
Data from the year 200-2002 are used to model and that from the year 2003 will be used to evaluate the performance of the model.

V. IMPLEMENTATION LOGIC

A. Data Preparation

The implementation begins by extracting the attributes RANK, SEX, CATEGORY, SECTOR, and BRANCH from the master database for the year 2000-2003 at the NODAL Centre. The database was not intended to be used for any purpose other maintaining records of students. Hence there were several inconsistencies in the database structure. By effective pruning the database was cleaned. A new table is created which reduces individual ranks to classes and makes the number of cases limited. All queries will belong to a fix set of known cases like:
RANK (1) SECTOR (U) SEX (M)
CATEGORY (GEN) BRANCH (A)
For every combination of these attributes, we calculate the corresponding placement chances from the history data by calculating probability of placement and storing in a separate data sheet which is used to build the model.
Probability (P) = Number Placed/ Total Number for this particular input combination
The chance is obtained by the following rules:
If P>=95 Chance='E'
If P>=75 && P<95 Chance='G'
If P>=50 && P<75 Chance='A';
Else Chance='P'; Where E, G, A, P stand for Excellent, Good, Average & Poor respectively.

B. Conversion For The Naive Bayes Classifier

The Explorer interface of WEKA has several panels that give access to the main components of the workbench. The Pre-process panel has facilities for importing data from a database, a CSV file, etc., and for pre-processing this data using a so-called filtering algorithm. These filters can be used to transform the data (e.g., turning numeric attributes into discrete ones) and make it possible to delete instances and attributes according to specific criteria. The Classify panel enables the user to apply classification and regression algorithms
(indiscriminately called classifiers in WEKA) to the resulting dataset, to estimate the accuracy of the resulting predictive model .The panel —select attributes" provides algorithms for identifying the most predictive attributes in a dataset.
WEKA input must be in ARFF format:
A relation name
– E.g., @relation student
A list of attribute definitions
– E.g., @attribute RANK numeric
@attribute SECTOR {U, R}
@attribute SEX {F, M}
@attribute CATEGORY {GEN, OBC, SC, ST}
@attribute BRANCH {A, B, C, D, E, F, G, H, I, J}
@attribute CHANCE {E, P, A, G}
– The last attribute is the class to be predicted
A list of data elements
@data
1, U, F, GEN, D, E

VI. PRINCIPLE OF DATA MINING BASED ON NEURAL NETWORK

Neural Network has the ability to realize pattern recognition and derive meaning from complicated or imprecise data that are too complex to be noticed by either humans or other computer techniques.
The data mining based on neural networks can generally be divided into 3 stages: data preparation, modeling and knowledge discovery.

A. Data Preparation

Data preparation is an important step in which the mined data is made suitable for processing. This involves cleaning data, data transformations, selecting subsets of records etc. Data selection means selecting data which are useful for the data mining purpose.

Data transformation or data expression is the process of converting the data into required format which is acceptable by data mining system. For ex: Symbolic data types are converted into numerical form understood by system. Also the data is scaled onto 0 to 1 or -1 to 1 scale which is acceptable by Neural Networks. Some sample mappings are shown in table I.

TABLE I. ATTRIBUTE VALUES MAPPED TO 0 TO 1 SCALE

ATTRIBUTE	RANGE	MAPPED TO
RANK	1 to 4000	0 to 1
SEX	1 to 2	0 to 1
CATEGORY	1 to 4	0 to 1
SECTOR	1 to 2	0 to 1
BRANCH	A to J	0 to 1
ACTIVITY	1 to 2	One of the four values 'E', 'G', 'A' and 'P'.

Table II shows a snippet of sample data used to train neural network.

TABLE II. SAMPLE OF DATA USED AS INPUT TO TRAIN THE NETWORK.

SEX	RESERVATION	LOCATION	RANK	BRANCH
0	0	1	0.72	0.47
1	0	1	0.72	0.47
0	1	0	0.59	0.33
0	0	1	0.4	0.66
0	1	0	0.72	0.47
1	1	1	0.27	0.38

The output data used for training is derived from ACTIVITY attribute. Instead of representing the output on 0 to 1 scale basis, we have used four fold classification that is, a 4 value code has been assigned with each record. A code value of 1000 represents a 'excellent' chances of getting a student placed, a code value of 0100, 0010, 0001 represents 'good', 'average' and 'poor' chances of placement of a student respectively.

The process of computing the placement chances of each possible attribute combination is same as that of the naïve Bayes classifier and its final assignment to codes are shown in table III.

MATLAB is used to model the neural network and scripts in MATLAB were used to model and test the neural network. The same work may be done using WEKA, but MATLAB gives more options and programmability for a neural network.

Range of Probability	Output Code	Chances of Placement
P >= 0.95	1000	Excellent
0.75 <= P < 0.95	0100	Good
0.50 <= P < 0.75	0010	Average
P < 0.50	0001	Poor

Table iii. Assigning Output Codes To Records

B. Modelling

This is the most important step in the data mining. A proper selection of algorithm is made on the basis of the required objective of the work.

One of the most popular Neural Network model is Perceptron, but this model is limited to Classification of Linearly Separable vectors[4]. The input data which is obtained from NTMIS, may contain variations resulting in Non-Linear data. For example a student with good rank may opt for a weak branch and still may have a placement. To deal with such inputs data cleaning alone is not sufficient. Therefore we go for multilayer perceptron network with supervised learning which gives back propagation Neural Network. A BP neural network reduces the error by propagating the error back to the network. Appropriate model of BP neural network is selected and repeatedly trained with the input data until the error reduces to a fairly low value. At the end of training we get a set of thresholds and weights which determines the architecture of the neural network. A back propagation neural network model is used consisting of three layers: input, hidden and output layers as shown in fig. 1. The number of input neurons is 5, which depends upon the number of the input attributes. The number of neurons used in hidden layer is 5; this number is obtained by value based on observations. A small number may not be able to solve a given problem and a large number of neurons increases the computation required. The number of neurons in output layer should equal the number of output code. Here we are performing 4 fold classifications, therefore four neurons are required at output layer. The total no. of records used for training is 4091 and total no. of records used for testing is 1063.

Transfer Functions- The transfer function used in the Hidden layer is Log- Sigmoid while that in the output layer is Pure Linear.

Learning- Training is done by using one of the Conjugate Gradient algorithm, Powell-Beale Restarts. This algorithm provides faster convergence by performing a search along the conjugate direction to determine the step size which minimizes the performance function along that line.

VII. TEST RESULTS

A. Naive Bayes Classifieer

Table Iv. Confusion Matrix (Student Data)

Confusion Matrix

		PREDICTED			
		E	P	A	G
ACTUAL	E	496	10	13	0
	P	60	97	12	1
	A	30	18	248	0
	G	34	19	22	3

For training, records of 2000-2002 are used and for testing the records of year 2003 are used. The predictions of the model for typical inputs from test set, whose actual data are already available for test comparisons, were compared with predicted values.

The results of the test are modeled as a confusion matrix as shown in the above diagram, as its this matrix that is usually used to describe test results in data mining type of Research works.

The confusion matrix obtained for the test data is as shown in table IV and the accuracy is computed as below:

AC = 844/1063 = 0.7938

To obtain more accuracy measures, we club the field Excellent and Good as positive and Average and Poor as negative. In this case we got an accuracy of **83.0%**. The modified Confusion matrix obtained is as shown in table V.

		Predicted	
		Negative	Positive
Actual	Negative	**365**	**101**
	Positive	**57**	**540**

TP = 0.90 *FP* = 0.22
TN = 0.78 *FN* = 0.09

B. Neural Networks

For simulation/evaluation we use the data of year 2003 obtained from NTMIS. The knowledge discovered is expressed in the form of confusion matrix in table VI.
Since the negative cases here are when the prediction was Poor /average and the corresponding observed values were Excellent/good and vice versa.

Table Vi. Confusion Matrix (Student Data)

Confusion Matrix

		PREDICTED			
		P	A	G	E
ACTUAL	P	31	4	1	91
	A	5	410	2	9
	G	1	1	6	12
	E	72	13	4	401

Therefore the accuracy is given by
AC = 848/1063 = 0.797
To obtain more accuracy measures, we club the field Excellent and Good as positive and Average and Poor as negative. Then the observed accuracy was 82.1 %. The modified Confusion matrix obtained is as shown in table VII.

Table Vii. Modified Confusion Matrix (Student Data)

		Predicted	
		Negative	Positive
Actual	Negative	**450**	**103**
	Positive	**87**	**423**

TP = 0.83 *FP* = 0.17
TN = 0.81 *FN* = 0.17

From these test results, now a methodology has to be found by which we can compare the model performances on this particular domain. The following section explains the techniques that are followed for comparing these two models.

VIII. Perfomane Comparison with other Models

The Kappa statistic values for neural networks and naive Bayes classifiers were found to be 0.56 and 0.69 respectively. The ROC values were 0.86 and 0.89 respectively.
For comparing model performances there exists a statistic which uses the classical hypothesis testing concept which may give a good measure on performance comparison,

$$P = \frac{|E1-E2|}{\sqrt{q(1-q)(2/n)}}$$

Where E1= error rate for model M1(Neural Network)
E2= Error rate for model M2 (Naive Bayes)
q= (E1+E2)/2
n=number of instances in test set
In our case E1=0.203, E2=0.206, n=1063,
q=0.2045
So applying these values in the formula P becomes 0.176, for the four variable output case.
According to classical hypothesis testing as **p<2** the difference in performance between the two models is not significant and is comparable or similar in their predictive capabilities with respect to this domain and test set.

IX. Conclusion

Choosing the right career is so important for any one's success. For that we may have to do a lot of history data analysis, experience based assessments etc. Nowadays technologies like data mining is there which uses concepts like Naïve Bayes prediction and neural networks, to make logical decisions. Hence this work is an attempt to demonstrate how technology can be used to take wise decisions for a prospective career. The methodology has been verified for its correctness and may be extended to cover any type of careers other than engineering branches. The work may be extended to analyze how data of other disciplines also may be modeled with these concepts.

X. References

1) SudheepElayidom.M, Sumam Mary Idikkula, Joseph Alexander-Applying data mining using statistical techniques for career selection, IJRTE ,ISSN 1797-9617, volume 1, number 1, May 2009.
2) Elizabeth Murray, Using Decision Trees to Understand Student Data, Proceedings of the 22nd International Conference on Machine Learning, Bonn, Germany, 2005.
3) U Fayyad, R Uthurusamy - From Data Mining to Knowledge Discovery in Databases , 1996.
4) Sreerama K. Murthy, Automatic Construction of Decision Trees from Data: A Multi-Disciplinary Survey, Data Mining and Knowledge Discovery, 345-389 1998.
5) L. Breiman, J. Friedman, R. Olshen, and C. Stone, Classification and Regression Trees, Chapter 3,Wadsworth Inc., 1984.
6) Kohavi R. and F. Provost, Editorial for the Special Issue on application of machine learning and the knowledge of discovery process, Machine Learning 30, 271-274, 1998.
7) M. Kubat, S. Matwin, Addressing the Curse of Imbalanced Training Sets: One-Sided Selection, Proceedings of the 14th International Conference on Machine Learning, 179-186, ICML'97.
8) Lewis D. D. & Gale W. A., A sequential algorithm for training text classifiers,3-12, in SIGIR'94.

9) SudheepElayidom.M, Sumam Mary Idikkula, Joseph Alexander ——Applying Data mining techniques for placement chance prediction". Proceedings of international conference on advances in computing, control and telecommunication technologies, India, December 2009.

Modeling And Implementing RFID Enabled Operating Environment For Patient Safety Enhancement

GJCST Classification
J.3 I.6.5

Chuan-Jun Su1

***Abstract*- Patient safety has become a growing concern in health care. The U.S. Institute of Medicine (IOM) report "To Err Is Human: Building a Safer Health System" in 1999 included estimations that medical error is the eighth leading cause of death in the United States and results in up to 100,000 deaths annually. However, many adverse events and errors occur in surgical practice. Within all kinds of surgical adverse events, wrong-side/wrong-site, wrong-procedure, and wrong-patient adverse events are the most devastating, unacceptable, and often result in litigation. Much literature claims that systems must be put in place to render it essentially impossible or at least extremely difficult for human error to cause harm to patients. Hence, this research aims to develop a prototype system based on active RFID that detects and prevents errors in the OR. To fully comprehend the operating room (OR) process, multiple rounds of on site discussions were conducted. IDEF0 models were subsequently constructed for identifying the opportunity of improvement and performing before-after analysis. Based on the analysis, the architecture of the proposed RFID-based OR system was developed. An on-site survey conducted subsequently for better understanding the hardware requirement will then be illustrated. Finally, an RFID-enhanced system based on both the proposed architecture and test results was developed for gaining better control and improving the safety level of the surgical operations.**

Keywords- Radio Frequency Identification (RFID), Patient Safety, Operating Room, IDEF0.

I. Introduction

The Operating Room (OR) briefing is a tool to enhance communication among the team members of operating room and improve patient safety [1], which is the most important and uncompromised issue for medical institutions. It has become a growing concern in health care. As expected, many adverse events and errors occur in surgical practice. Taking the correct patient into the correct OR and executing correct procedures by correct medical staff have become widely understood as the fundamental infrastructure of safe patient care to avoid adverse events in the operating room. Hence, there are four critical requirements to give the right treatment to the right patient:

i. Correct patient
ii. Correct OR
iii. Correct medical staff
iv. Correct operations

Whether wrong patient, wrong location, wrong medical staff or wrong operation event has resulted in injury or didn't raise actual harm, those kind of events cause anxiety for patients and staff, disrupt the smooth flow of patients through the OR suite, and increase the probability of medical errors. Hence, this research aims to develop a prototype system based on RFID that detects and prevents errors in the OR. The system provides hospitals to correctly identify surgical patients and track their operations to ensure they get the correct operations at the right time.

II. Literature Review

Patient safety means minimizing harm to patients arising from medical treatment [2] and is becoming a growing concern in health care. There are many factors involved in patient safety today, and there are many factors to consider when improving the safety process [3]. Recent attention to this topic stems from several high-profile medical errors and several Institute of Medicine (IOM) reports which quantified the problem, created standardized definitions, and charged the healthcare community to develop improved hospital operating systems [4, 5].

The operating room (OR) is one of the most complex work environments in health care. Compared with other hospital settings, errors in the operating room can be particularly catastrophic and, in some cases, can result in high-profile consequences for a surgeon and an institution. In addition, the high rate of adverse events in surgery is incessantly demonstrated. According to a sentinel event alert issued Dec 5, 2001, by the Joint Commission on Accreditation of Healthcare Organizations (JCAHO), "fifty-eight percent of the cases occurred in either a hospital-based ambulatory surgery unit or freestanding ambulatory setting, with 29 percent occurring in the inpatient operating room and 13 percent in other inpatient sites such as the Emergency Department or ICU. Seventy-six percent involved surgery on the wrong body part or site; 13 percent involved surgery on the wrong body part or site; and 11 percent involved the wrong surgical procedure" [6]. After the astonishing report published by JCAHO Gawande et al., in 2003, analyzed errors reported by surgeons at three teaching hospitals and found that seventy-seven percent involved injuries related to an operation or other invasive intervention (visceral injuries, bleeding, and wound infection/dehiscence were the most common subtypes), 13% involved unnecessary or inappropriate procedures, and 10% involved unnecessary advancement of disease. In addition, two thirds of the

Chuan-Jun Su[1]
[1]Department of Industrial Engineering & Management, Yuan Ze University 135, Far-East Rd., Chung-Li, Taiwan, ROC
E-mail: iecjsu@saturn.yzu.edu.tw

incidents involved errors during the intra-operative phase of surgical care, 27% during pre-operative management, and 22% during post-operative management [7]. In other words, no matter how well-trained a medical staff is, he or she could still make mistakes.

Within all kinds of surgical adverse events, wrong-side/wrong-site, wrong-procedure, and wrong-patient adverse events (WSPEs) are the most devastating, unacceptable, and often result in litigation. However, an estimate of 1300 to 2700 WSPEs per year based on the available databases, extensive review of the literature, and discussion with regulators in the United States seems likely [8]. A variety of studies have demonstrated that the rates of adverse events associated with surgery are substantial. Of course, surgery inherently carries risk, and only 17 per cent of these adverse events were judged to be preventable [9].

Nonetheless, this important proportion of surgical adverse events is preventable given what is known today. Systems must be put in place to render it essentially impossible or at least extremely difficult for human error to cause harm to patients. With the introduction of new approaches many other complications that are not associated with an obvious error may be preventable in the future.

III. Requirement Study

As mentioned in the previous chapters, the high rate of sentinel events in surgery has been incessantly demonstrated. Among all sentinel events, performing a procedure on the wrong site or the wrong patient is mostly preventable and should never happen. Much literature claims that systems must be put in place to render it essentially impossible or at least extremely difficult for human error to cause harm to patients. Among a lot of novel technologies, RFID is an enabling technology that is generally considered to improve patient safety and savings in hospital. This technology has been applied for many fields but few applications specified to OR. Moreover, little literature analyzed from the business processes‘ point of view to reap the benefits of RFID but focus on an object or an individual. Therefore, we proposed using RFID from the processes‘ point of view to detect errors that may lead to wrong site or the wrong patient surgery.

Before an RFID implantation, opportunity survey based on business process analysis is necessary. The survey consists of the following phases:

Expert interview and site survey- Expert interview and site survey have conducted to comprehend the process in OR.

Existing OR Process: Based on the result of expert interview and site survey, we described the existing OR process.

IDEF0 modeling-*:* IDEF0 modeling technique is adopted to (1) build the OR as-is model based on the result of previous step, (2) analyze the activities in the previous OR process.

RFID-based OR Process: Based on the results of previous steps, we described an RFID-based state for the process.

A. Expert Interview and Site Survey

Before an RFID implantation, opportunity survey based on business process analysis is necessary. This experience can also be applied in a hospital. Surgeons, nurses and anesthetists are the experts who know what happens behind the closed doors of the operating department.

They clearly know their job functions and workflow during surgery. For this reason, we undertook several expert interviews with the medical staff worked in an operating department of a regional teaching hospital in Taoyuan to comprehend the operative process. Except expert interviews, we also perform site survey to observe the activities in OR. The activities during on-site survey include observation of nursing work, review related forms and face-to-face meeting with nurses to capture the entire workflow in OR. The results of expert interviews and on-site survey are described in the following sections.

B. Expert Interview And Site Survey

The scope of OR process for an individual patient which we describe as follows begins with the surgical patient‘s arrival at the OR suite and ends when the patient leaves the OR suite. The details of the OR process is shown below:

Admission into operating suite- Base on operations schedule, the transporter brings scheduled surgical patient from ward to the operating suite, along with his/her medical record and related document. Upon the patient‘s arrival in the holding area of operating suite, the holding area nurse orally identifies patient by matching the replies from the patient about the name, ID card number, type of surgery with medical record, etc. After the confirmation, the nurse reviews the document accompanying the patient to check whether the operation related forms such as operative consent form has been completely filled in or not. The patient‘s national health insurance card is then received by the nurse. After a series of admission procedure, the nurse logged on to the hospital information system to change the patient‘s status. At the same time, the patient‘s status information ―Waiting for surgery‖ is displayed in the screen located in the waiting area to reduce anxiety patient‘s family members.

Admission into operating room- The surgical patient stays in the pre-operative holding area until the OR is ready. However, before a circulating nurse takes a patient into scheduled OR, the nurse verbally identifies the patient again and changes the patient‘s status from ―Waiting for surgery‖ to ―In surgery‖.

Beginning of anesthesia- The anesthetist verbally confirms the patient‘s identification, type of surgery and part of surgery before anesthesia. If the information is correct, the anesthetist signs in the nursing records of patients' operations. Although the doctor has already determined the anesthetics and methods of delivery before surgery, the anesthetist can make the final decision depending on the patient‘s specific condition at that point of time. The anesthetist verbally asks the patient‘s information such as the history of allergy, family history in the OR to decide which kind of anesthetics he or she should use. Besides, anesthesia staff also reviews the anesthesia information and laboratory test data in medical record to assist determining

the way to induce anesthesia is suitable for the patient or not.

Surgery- A surgeon also has to confirm the patient‘s identification before surgery. However, the surgical patient is usually covered with surgical drapes and has been anesthetized when a surgeon enters the OR. The surgeon can only justify the patient by medical record or pictures such as X-ray pictures. In the case that the patient was not covered with surgical drapes, the surgeon verifies patient by face. Because surgeons often meet surgical patients before surgery, the surgeons consider that they can distinguish a right patient from a wrong patient by their memory. Before performing an operation, the surgeon refers to the patient‘s medical record to make sure the surgical procedure and the site of operation. During surgery, surgeon can refer to medical record for the anesthesia information, patient information, results from laboratory information system if necessary. In the operating room information about the location of equipment had direct and sometimes critical implications for patients‘ clinical outcomes [8].

Admission to recovery- After surgery the patient is takes into recovery rooms to wait for "awakening" from anesthesia. At the same time, the nurse in the recovery room changes the patient‘s status to” In recovery”.

Discharge from operating suite: After a patient becomes conscious, the transporter takes the patient back to his/her ward and change patient‘s status information to —Return ward”.

C. *IDEF0 Modeling*

After in-depth understanding of the current process in OR an IDEF0 model was developed. The IDEF0 model is used to help organizing the analysis of OR system and to promote good communication between the analyst and the medical staff. For better understanding of the sequence, we chose sequential form of breakdown which decomposes the parent activity by a sequence of sub-activities to build our system. However, the structure imposed by the IDEF0 methodology naturally creates a set of questions that must be asked and answered about each function and its sub-functions. The answers to these questions provide important information concerning how known human fallibilities which may lead to errors. Thus, we can clarify many activities during model development stage by discussing with medical staff.

In this section, first, we built the —as-is” model to define activities and functions in OR. The definition of —as-is” is a description of the current situation in terms of the work processes. With sufficient information regarding the as-is operation, analyzing current process and building a new system become easier. Second, we examined the model, found out the operations which probably threaten to patient safety or make medical staff ineffective and analyzed the opportunity for introducing RFID to solve the problem.

D. *Building “as-is” Model*

After expert interviews and on-site survey, the OR as-is model was constructed. The purpose of this model is to find out the systemic vulnerabilities that may lead to human error and the opportunities that can improve medical staff‘s operations by introducing RFID technology. The model was developed from the OR medical staff‘s viewpoint.

Fig.1 depicts the top-level function of the IDEF0 model, —perform surgery”. The activity called —performing surgery” is broadly defined as all activities during per-, intra- and post operations. Fig.2 shows the top-level function decomposed into six more specific functions, representing the surgery process in more detail. Note that the general inputs, outputs, controls, and mechanisms from Fig.1 are also decomposed, illustrating the progressive exposure of detail that is a feature of the IDEF0 methodology; Fig.3 shows the decomposition of the activity —patient check-in” at an even higher level of detail. In the same way, Fig.4, 5 and 6 shows the decomposition of the other activities in detail.

E. *As-is Model Analysis and RFID Solutions*

In the operating room information about the location of equipment had direct and sometimes critical implications for patients‘ clinical outcomes [10]. After building the OR as-is model and numbers of meeting with related medical personnel, we discovered that there were a lot of systemic vulnerabilities that may lead to human error. In real situations, not every OR member completely follows the Standard Operating Procedures (SOP). If some accidental situations happen, wrong patient, wrong site/side surgery, unsuitable anesthesia or wrong OR event may occur. Based on the as-is model, the possible human errors and inefficient operations are listed as follows:

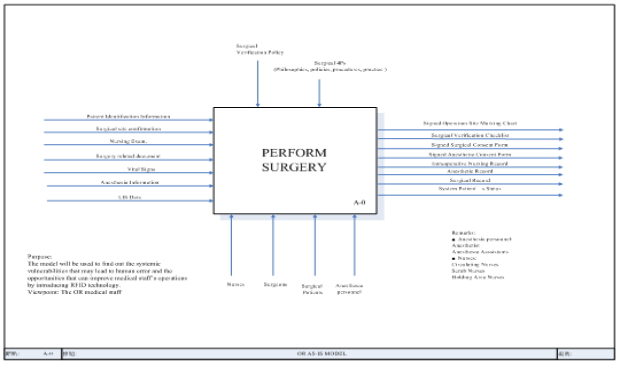


Fig.1. Or As-Is Model

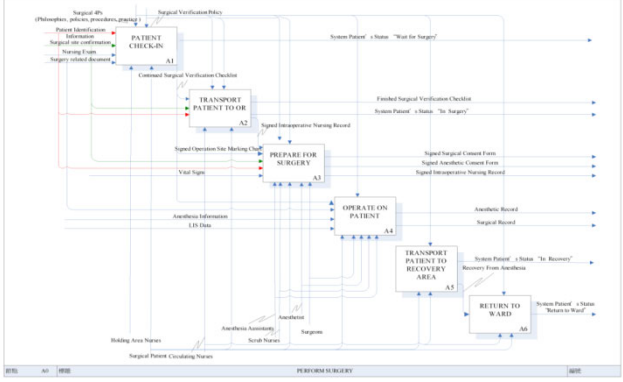


Fig.2. Idef0 Diagram “Perform Surgery”

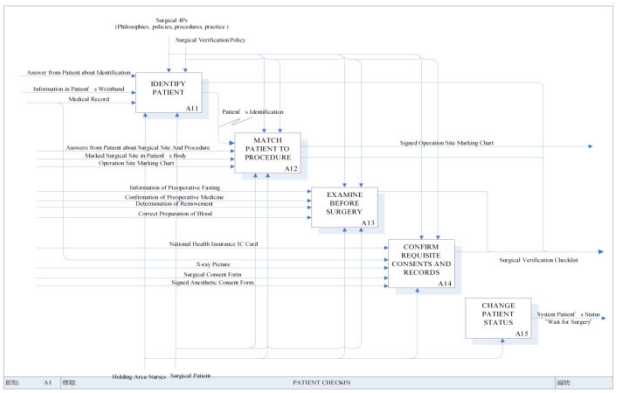


Fig.3. Idef0 Diagram "Patient Check-In"

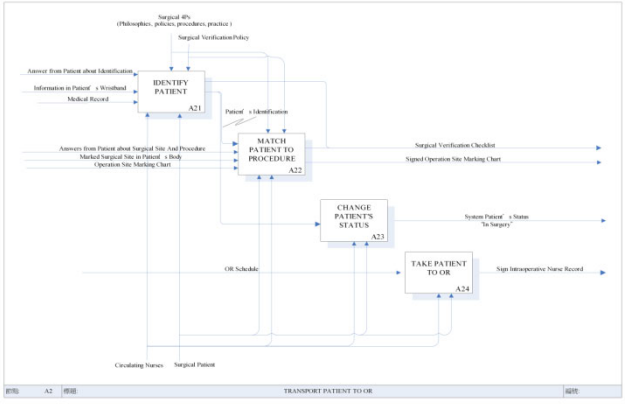


Fig.4. Idef0 Diagram "Transport Patient To Or"

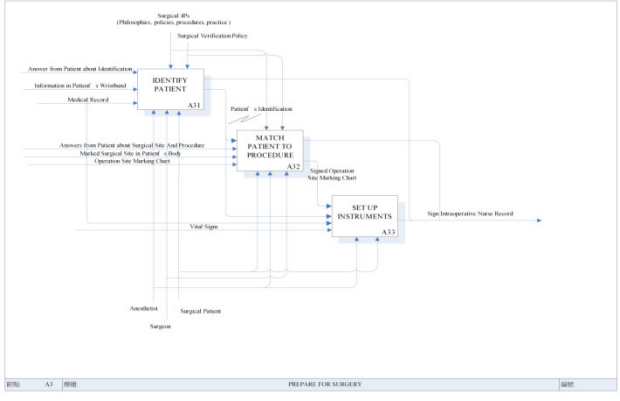


Fig.5. Idef0 Diagram "Prepare For Surgery"

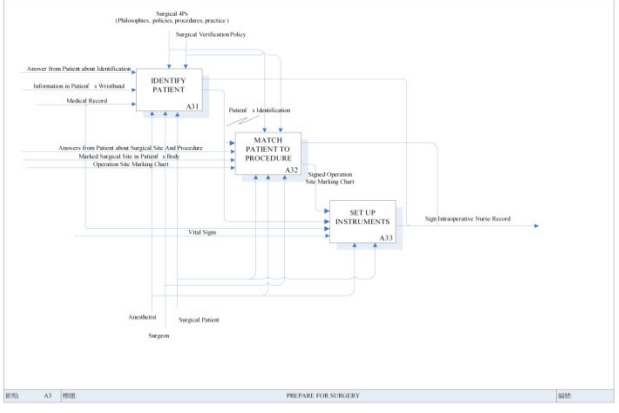


Fig.6. Idef0 Diagram "Operate On The Patient"

Misidentification Of Patients- According to the as-is model, the nurse verifies the patient's identity by asking the patient his/her full name and checks it with both the patient's identification bracelet and medical record. If frontline healthcare staff did not successfully verify a person's identity, wrong patient may be taken into OR. In other words, if an exceptional case happens, the manual process is probably permitted errors to cause wrong patient surgery. Failure to correctly identify patients constitutes one of the most serious risks to patient safety; however, in the OR, it can even cost a life.

Entering the wrong OR- In general, a teaching medical center has numbers of contiguous ORs, each performing two to three cases daily. Moreover, OR schedule varies frequently. Hence, there are many opportunities for surgeons or patients to enter the wrong ORs. If a patient or a surgeon enters a wrong OR without reconfirm, the event of wrong surgery could happen.

Inducing unsuitable anesthesia- An anesthetist has to make the final decision regarding the anesthetics and methods of delivery at the time of surgery. Therefore, providing sufficient information to the anesthetist is very important. In general, the critical information that aids an anesthetist making decision must be obtained from both ways: verbally asking the patient and reviewing some data described in the medical record. The questions now arise: first, some critical information even was not recorded in a patient's medical record. History of allergy and family history, for example, only can be received by asking patients. If the patient is not conscious or too elderly to answer the questions from medical staff, it would cause the patient to be exposed to danger. Second, looking for the unfiltered data on the paper is inefficient to anesthesia staff.

Performing wrong operations- Wrong operation mentioned here includes performing wrong procedures on a patient or perform a surgery on the wrong site/side of a patient's body. Based on the as-is model, the doctor conforms patient's operation-related information by checking the patient's medical record and surgery-site chart that describes the surgical site/side of a patient. However, because it is not convenient to find out disperse data by looking up the paper-based medical records, doctors sometimes depend on their memory for patients' condition without double-check. In addition, when a surgeon substitutes for another surgeon to perform a surgery or a surgeon operates on more than one patient at the same time, the wrong operation may be performed.

Inefficiently updating patient's states-Based on the as-is model, nurses have to manually update a patient's status information (in surgery, in recovery, etc.) in the hospital information system when the patient's status is changed. However, the essence of the OR nurses is to provide care and support to patients before, during, and after surgery. This kind of unrelated activity distracts medical staff and obstructs medical professionals providing better patient care.

Table1. Defects in the as-is Model

Events That May Threaten Patient's Safety	Node Number	Accidental Case (Do not follow the SOP)
- Wrong Patient	A11, A21, A31	A surgeon directly brought a patient to an OR without reconfirmation. Just call the first name of a patient with Mr. or Ms. to verify the patient. (A21) A nurse misidentifies a patient because of fragmented communication between the nurse and the patient.
- Wrong OR	A24	Mistakenly bring a patient into a wrong OR.
- Wrong Procedure	A42	Passive and inconvenient confirmation processes do not encourage medical staff to reconfirm patient's information. A substitute surgeon is not familiar with the surgical patient. A surgeon operates on more than one patient simultaneously.
- Wrong Anesthesia	A41	A patient who is not conscious or too elderly to answer the questions from medical staff causes that some critical information can not be obtained. It is inconvenient to look for dispersed and unfiltered information.
- Inefficient Operations	A15, A23, A5, A6	Wasting time to update patients' status information in the computer obstructs medical professionals providing better patient care.

As mentioned above, not every OR member completely follows the SOPs in real situations. Therefore, if some accidental situations happen, wrong patient, wrong site/side surgery, unsuitable anesthesia or wrong OR event may occur. Table1 shows potential cases that may threaten patient's safety in the as-is model.

Human error is inevitable and unavoidable. However, most preventable adverse events are not simply the result of human error but are due to defective systems that allow errors to occur or go undetected. Therefore, we propose an RFID-based OR system that can reinforce SOP and prevent potential errors for achieving the ultimate objective of improving patient safety in OR. The proposed RFID-based OR system is expected to complement current human-based operations in the following ways:

Patient Identification- Improving the accuracy of patient identification is one of JCAHO 2006-2007 patient safety goals which suggest using active communication technique to conduct final verification process and using at least two patient identifiers. However, either suggestion was not adopted in the "as-is" process. The problem can be solved by introducing RFID that can automatically identify patients to complement current human-based verification.

Surgical site verification-To decrease the incidence of wrong site/side surgery, we have developed a digital chart of surgery site marking which can be displayed on the LCD monitor of OR. By integrating RFID with hospital information system (HIS), the digital chart and some critical information of the patient are automatically shown on the monitor when the patient is brought to the scheduled OR.

OR verification-If a patient is brought to an unscheduled OR, the system will create a warning on the monitor. Thus, the RFID-based system checks the wrong-location event to prevent the potential wrong surgery.

Patient status update-The activity of updating patient's status distracts medical staff from surgery related tasks. In the developed RFID-based system, the status will be updated automatically by integrating RFID with the back-end HIS.

F. RFID-enabled OR Process

After analyzing as-is model, we have developed an RFID-based prototyping system that detects and prevents errors in the OR. The system provides hospitals to 1) correctly identify surgical patients, 2) track the ORs in which patients and medical staff enter, and 3) furnish critical information to ensure patients get the correct operations at the right time and place. The "to-be" RFID-based OR process derived from analyzing the "as-is" model is illustrated as follows:

Admission into the operating suite- When a surgical patient is scheduled to be operated upon, the nurse in the ward assigns the patient an RFID-embedded wristband encoded with a unique ID. When the surgical patient is brought to the holding area in the OR suite, an RFID reader automatically verifies the identity of the patient. If the details of the patient and operation schedule match, the

monitor in the holding area displays some brief information about the patient. Simultaneously, the system automatically changes the patient's status information to "waiting for surgery" in the database. Then, the patient's status information can be displayed on the screen located in the waiting area that it can helps in reducing the anxiety of his/her family members.

Admission into operating room- The RFID readers in the OR automatically capture the information on the patient's tag to identify him/her upon entering the OR. If the details of the patient and the OR into which he/she has entered match, the screen in the OR displays the patient's information, including his/her name, age, gender, laboratory test data, digital surgery-site chart, and scheduled procedure, by associating the tag's ID number with the patient records stored in the hospital information system. However, if an unscheduled patient enters the OR, the system can alert the medical staff to take the necessary measures in the OR. Thus, unfamiliar faces can be checked with assurance, thereby decreasing the probability of performing the procedure on a wrong patient. Subsequently, the system automatically changes the patient's status information to "in surgery" after confirming that the correct patient has entered the correct OR. In addition, because the time is automatically recorded, the medical staff does not have to record the time manually. Thus, unfamiliar faces can be checked with assurance, thereby decreasing the probability of performing the procedure on a wrong patient. Subsequently, the system automatically changes the patient's status information to "in surgery" after confirming that the correct patient has entered the correct OR. In addition, because the time is automatically recorded, the medical staff does not have to record the time manually.

Initiation of anesthesia: When a tagged anesthetist or nurse enters the room, the system will also ensure that the person has entered the right room, thus preventing medical staff from rushing into the wrong OR and administering inappropriate medical treatment. In addition, information such as the history of allergy, family history, and laboratory test data, which aids the anesthetic team in determining the appropriate anesthetics and method of administration, is displayed on the monitor when the anesthetic team enters the OR. Furthermore, any abnormal values will be marked in red to bring them to the medical staff's attention.

Surgery- In the same manner, when a surgeon enters the room, the system will also check whether the person has been assigned to the room, in order to prevent doctors from entering the wrong OR and performing surgery on the wrong person. If the patient has not yet been covered with surgical drapes, the surgeon can reconfirm the patient's identity by matching the patient with a photograph displayed on the screen. In addition, the surgeon is also provided with the patient's critical information on the digital display, rather than having to search for the pertinent information in the documented reports. Thus, bringing all this information together not only saves time but also increases patient safety. The surgeons are encouraged to confirm the patient records through the centralized data displayed on the monitor because of this increased accessibility. In the absence of such a system, doctors would generally depend on their memory, without performing any reconfirmation due to the inconvenience involved.

Transfer to recovery- After surgery, patients are transferred to the recovery area. Here, the RFID readers detect the patient's RFID tags and the system automatically changes the patient's status to "in recovery."

Discharge from operating suite: After a patient recovers from anesthesia, the transporter brought the surgical patient leaving recovery room and returning to ward. At this time, the system automatically updates the patient's status information to "returned to ward."

IV. The Architecture of RFID-Based OR System

Beside on the result of the previous section, we proposed an architecture of the RFID-based OR system. The architecture of the RFID-based OR system consists of 1) physically distributed RFID readers, tags, 2) an RFID server that processes the data from the readers, 3) several client PCs that run different hospital applications, and 4) the hospital information system (HIS) that plays the same role as that of an ERP system in enterprise-level architectures. The RFID server contains a backend database and software called the concentrator that receives the data from RFID readers when the tags are detected. Then, the concentrator checks this data for errors and stores it on an operational database. In our architecture, the concentrator also communicates with other software that implement the application's business logic on client PCs. In addition, the RFID server is connected to the HIS to extract hospital data through an intranet. The architecture of this system is illustrated in Fig.7.

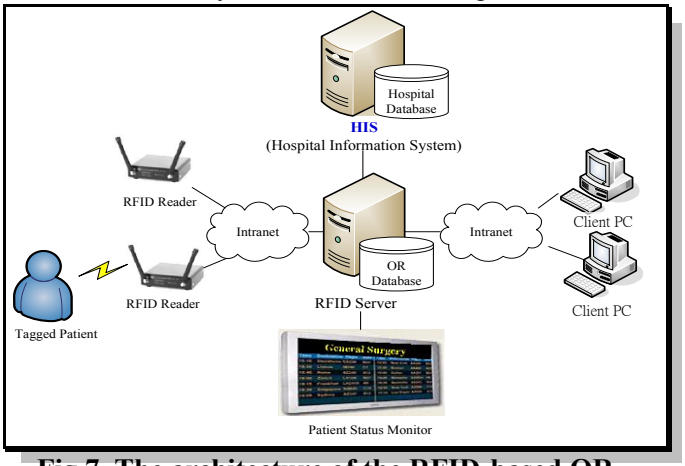


Fig.7. The architecture of the RFID-based OR System.

In summary, there are five primary components in the RFID-based OR system: (1) Reader, (2) tag, (3) RFID server, (4) client PCs that manage the client-side applications, and (5) the HIS.

V. Implementation

A. RFID Hardware Test and Deployment

A "Site Survey" is essential for real RFID implementation.

The actual RF coverage, number of readers, their placement and configuration, and system accuracy are greatly dependent on environmental factors.

In order to increase the feasibility of the hardware in a clinical environment, we deployed our hardware in an OR suite in a regional teaching hospital in Taoyuan to determine the optimum hardware deployment that fits our test scenario. The test environment was set up in three regions of an OR suite including the holding area, OR-5, and the recovery area. After a series of tests, the result each test scenarios is show as Table 2 and the RF coverage as shown in Fig.8.

Table2. Results of Testing

Test Scenario	Goals	Results
Holding area	The reader located in the holding area steadily detects the investigator‘s arrival.	○
	Do not detect a passerby out of the OR suite	○
	Avoid detecting other patients who passed the holding area on the corridor in the OR suite	X
OR 5	The tagged investigator must be detected by the reader located in OR-5 when he/she entered OR-5	○
	Do not detect other patients who passed OR-5 on their way to the correct locations	○
	RFID signals should not be picked up through the walls of a room adjacent to OR-5	○
Recovery area	The tagged investigator must be detected by the reader located in the recovery area while he/she entered the area.	○
	The reader located in the recovery area cannot detect other patients who passed the recovery area on their way to their scheduled ORs.	○
	The reader also can‘t detect a patient who walked on the corridor outside the OR suite.	X

Although part of testing results does not meet our requirement, we can use a system to filter certain unnecessary signals. The solution can be summarized as shown in Table 3.

Table3. Solution of Redundant RF Coverage

Location of readers	Solution of exceeded RF signals
Holding Area (Exception 1)	If a patient who has already been detected by the reader located in the holding area is detected again, the application system executed in the holding area will omit the event to avoid the patient check-in twice. Therefore, the RF coverage can extend beyond the zone of the holding area because the detection of a passerby walking on the corridor in the OR suite will not be processed again.
Recovery Area (Exception 2)	The application system executed in the recovery area supposes a patient who was not detected by any other readers located in the OR suite was a passerby. Hence, if a patient who do not have to undergo an operation passed through the recovery area out of the OR suite and is detected by the reader located in the recovery area, the system will filter the event to avoid mistakenly execute the application.

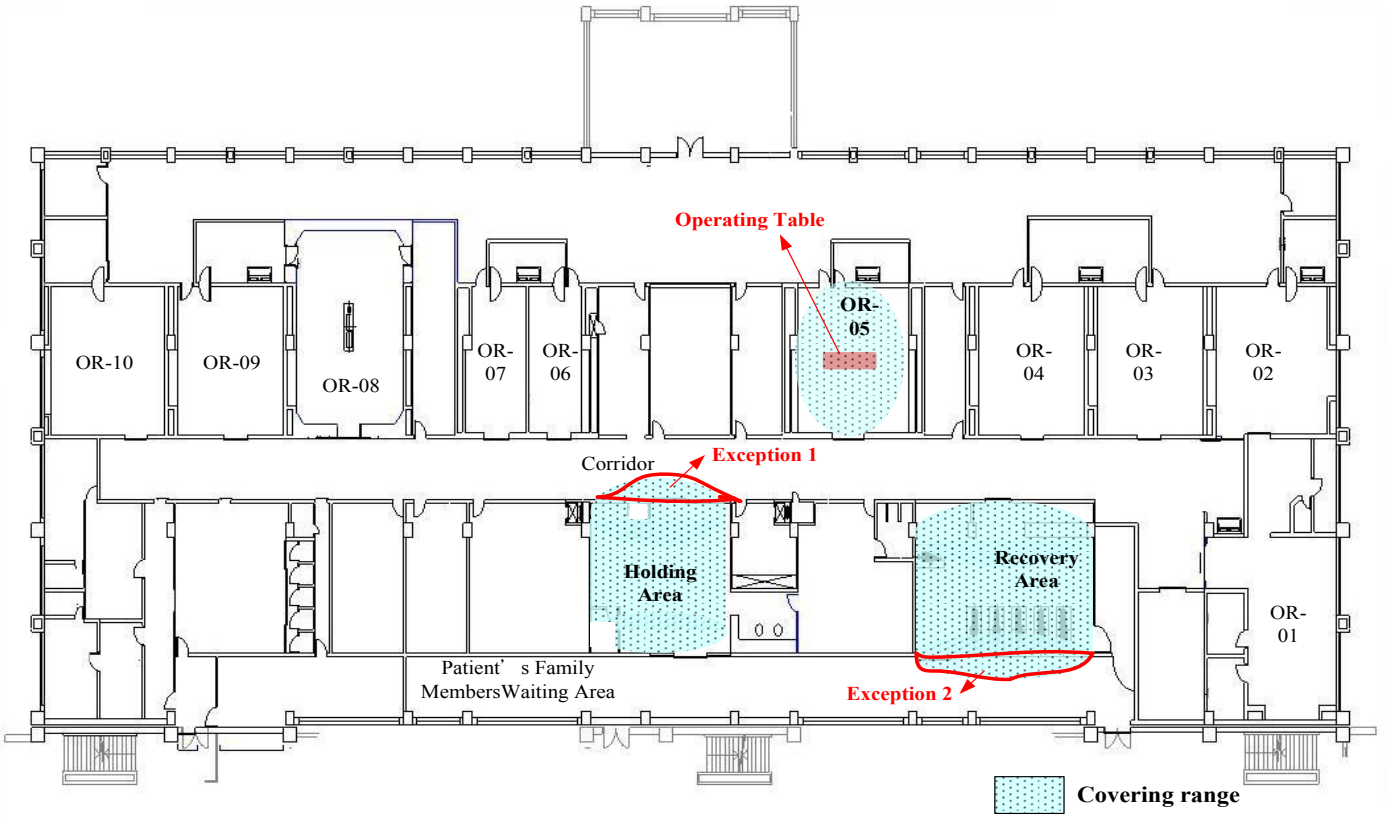


Fig.8. RF coverage in the clinical environment

B. *Two Different Types Of Scenarios*

After RFID hardware test and deployment, we have proposed two different types of scenarios: normal and accidental situations. The scenario of normal situations describes the general process during surgery; the scenarios of accidental situations show the unexpected events that may threaten patient‘s safety. Those scenarios was demonstrated by our system to describe how the RFID-based OR system prevents errors and improve patient safety.

Scenario 1: Normal situations- Miss Wang (Shiau-ching Wang, a pseudonym) is admitted to the obstetrics and gynecology department of a teaching hospital for surgery. The bed number of Miss Wang is W06-1. The doctor in charge of Miss Wang logged on to the admission application system to enter some surgical information about Miss Wang and to draw the operation site in detail.

Based on the operations schedule, the holding area nurse telephoned the obstetrics and gynecology floor, identified herself by name, and asked for Miss Wang to be prepared for the operation. After the phone call, the nurse checked Miss Wang‘s medical records, logged on to the admission application system, and assigned an RFID wristband to her. After a while, the transporter arrived at the obstetrics and gynecology department and brought Miss Wang from the ward to the operating suite, along with her medical records and related documents.

When Miss Wang entered the holding area of the OR suite, the reader grabbed the tag‘s ID stored in the RFID wristband. Subsequently, the monitor in the holding area displayed brief information about Miss Wang. The screen located in the waiting area outside the OR suite simultaneously displayed the patient‘s status information as ―waiting for surgery” to reduce the anxiety of the patient‘s family members. The holding area nurse completed the admission procedure for Miss Wang, after which Miss Wang remained in the holding area waiting for surgery.

After a while, the OR was ready for surgery. A circulating nurse took Miss Wang into the scheduled OR-5. While Miss Wang was brought to OR-5, the reader detected Miss Wang‘s RFID wristband and the monitor located in that OR displayed Miss Wang‘s personal and surgery-related information. At the same time, the screen located in the waiting area updated the patient‘s status information as ―in surgery.”

An anesthetist, Dr. Chen, entered the OR, and checked the information that would help him make the final decision regarding the anesthetics and methods of delivery. The surgeon, Dr. Lin, entered OR-5 and reviewed the information displayed on the monitor to reconfirm the patient‘s surgical procedures. Following these checks, the surgery was performed.

After surgery, Miss Wang was brought to the recovery area. The readers in the recovery area detected the RFID wristband worn by Miss Wang. Hence, the monitor in the recovery area displayed brief information about Miss Wang. At the same time, the screen located in the waiting area updated the patient‘s status information as ―in recovery.”

After Miss Wang recovered from the anesthesia, the transporter brought her from the recovery room to her ward. While Miss Wang has leaved the recovery room, the screen located in the waiting area updated the patient‘s status information as ―returned ward.”

Demonstration of System Usage Scenario 1-

After the decision to perform an operation on Miss Wang, the doctor in charge of Miss Wang logged on to the admission application system (See Fig.9), pressed the button marked ―operation site mark,” and inputted two identification parameters—the number on the identification card and the bed number—in order to provide some surgical

information about Miss Wang and to draw the operation site in detail (See Fig.10 and Fig.11).

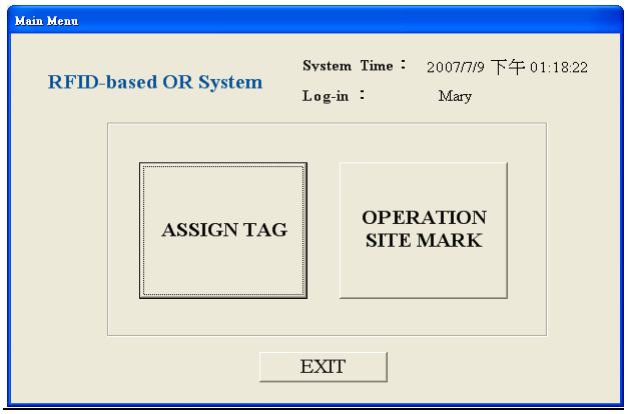


Fig.9. Main menu of the admission application system

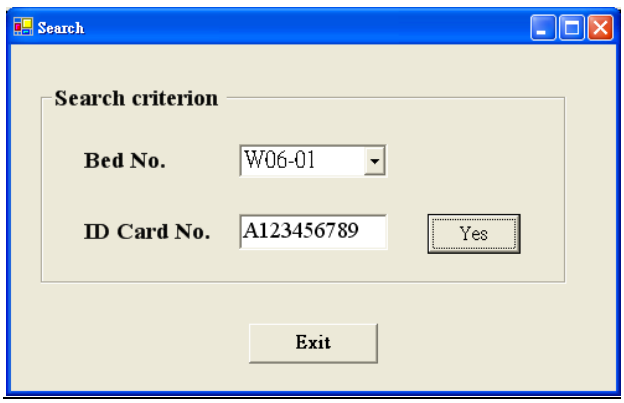


Fig.10. Searching Interface

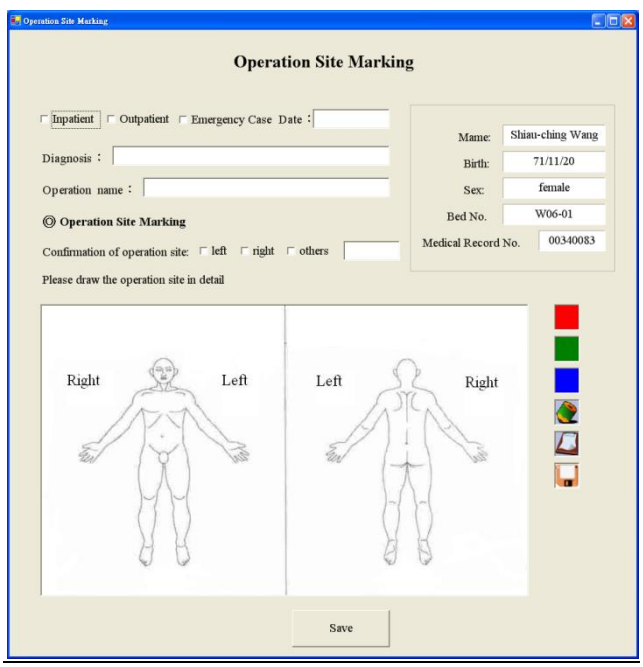


Fig.11. Operation site marking

After the phone call from the OR, the nurse working in ward 6 checked Miss Wang‘s medical records, logged on to the admission application system, and assigned an RFID wristband to Miss Wang (See Fig.12). Miss Wang entered two identification parameters—the number on the identification card and the bed number—to retrieve Miss Wang‘s information for double-checking. At the same time, the system automatically checked the operation schedule to confirm Miss Wang‘s operation. After ensuring that Miss Wang was the correct patient whose name had listed in the operation schedule, the system associated the ID stored in the RFID wristband with the patient‘s identification in the database. After this computerized process, the nurse assigned the RFID wristband to Miss Wang.

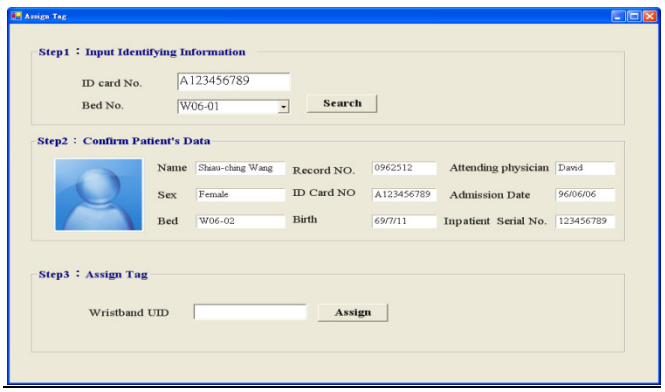


Fig.12. Assign tag

On Miss Wang‘s arrival in the holding area of the operating suite, the preoperative application associated the tag‘s ID propagated from the middleware with the patient‘s identification and checked the patient‘s identification with the operation schedule. After confirming Miss Wang‘s identification, the monitor in the holding area displayed brief information about Miss Wang. Subsequently, the system automatically changed the patient‘s status to ―waiting for surgery” (See Fig.13).

Fig.13. User interface displayed in the holding area

While Miss Wang was brought to OR-5, the operative application updated Miss Wang's status to "in surgery" to inform her family about the progress of the operation. At the same time, the system automatically executed the user interface and displayed some information about Miss Wang including personal information, surgical site, diagnosis, operation description, laboratory test data, etc. to avoid fragmented communication and dispersed information (See Fig.14 and Fig.15).

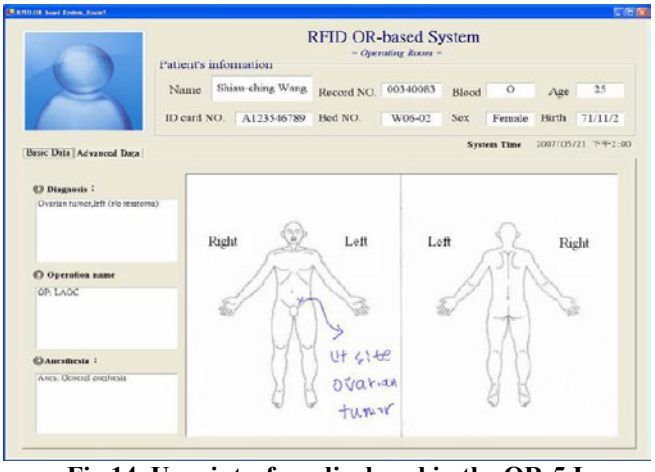


Fig.14. User interface displayed in the OR-5 I

The information displayed on the screen can be separated into three parts: basic patient information (helps identify the patient), operative information (ensures correct procedures), and advanced information (supports anesthesia decision). Basic patient information includes the patient's picture, name, ID card number, medical record number, bed number, blood type, sex, age, and birthday. Operative information includes the diagnosis, operation description, and the anesthesia administered by the doctor. Advanced information contains a history of allergy, information of diseases in the family, chronic prescription medicines, and laboratory test data. In addition, abnormal values of certain parameters obtained from laboratory results are marked in red to serve as a warning.

Fig.15. User interface displayed in the OR-5 II

Miss Wang was taken to the recovery room post surgery. On her arrival in the recovery room, the reader located in that room detected her RFID wristband. Subsequently, the post-operative application displayed brief information about Miss Wang on the monitor, changed Miss Wang's status to "in recovery," and recorded the arrival time in the database (See Fig.16).

After Miss Wang recovered from anesthesia in the recovery room, the transporter brought Miss Wang from the recovery room to her ward. Because Miss Wang had leaved the recovery room, the readers were unable to receive the signal from the patient's tag. Therefore, Miss Wang's information was no longer displayed on the monitor and her status was updated to "return to ward." In addition, the screen located in the waiting area also updated the patient's status information to "return to ward."

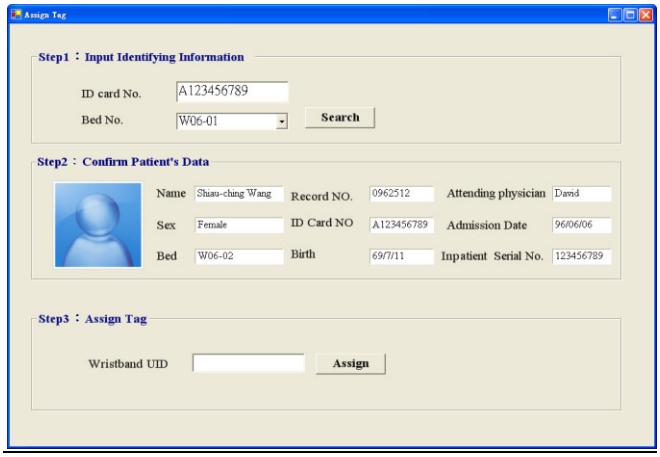


Fig.16. User interface displayed in recovery area

Scenario 2: Accidental Situations-As mentioned in chapter3, not every OR member completely follows the SOPs in real situations. Therefore, of some accidental situations happen, wrong-patient, wrong-site/side surgery, unsuitable anesthesia or wrong-OR event may occur. The following scenarios show probable situations that may threaten patient's safety in the as-is model.

Wrong OR Event-Based on the abovementioned scenario (Scenario 1), another patient who was assigned to OR-1 was mistakenly brought to OR-5 by a careless circulating nurse before Ms. Wang was brought to the room.

Wrong Side Surgery Event- Based on the abovementioned scenario (Scenario 1), Ms. Wang's attending physician unexpectedly cannot perform the operation due to some reason. Another doctor, Dr. Chen, replaced her attending physician; this doctor was not familiar with the patient's condition. Furthermore, he memorized the wrong side of the operation site and did not check the patient's report.

Wrong Patient Event-Another patient, a woman with a similar name (Shiau-chin Wang, a pseudonym) was also admitted to the obstetrics and gynecology department; however, her operation was planned for the next day. The holding area nurse telephoned the obstetrics and gynecology floor, identified herself by name, and asked for "patient Shiau-chin Wang" (giving no other identifying information).

The nurse on the other end listened to the information, but did not reconfirm; the nurse mistook —Shiau-ching Wang" for —Shia-chin Wang" informed the actual patient (Shiau-chin Wang) that she would have to undergo an operation that same day and prepared her reports. The patient, although feeling slightly confused, assumed that the operation was advanced by one day. After a while, the transporter arrived at the obstetrics and gynecology department and asked for —patient Ms. Wang" (giving no other identifying information). Consequently, the wrong patient was brought to the OR suite.

Demonstration of System Usage Scenario 2

Wrong OR Event- While the wrong patient was brought to OR-5, the reader located in OR-5 detected the patient's identity and checked it with the OR schedule. Since the patient was not assigned to OR-5, the system displayed a warning message bringing this fact to the nurse's attention (See Fig.17).

Fig.17. Warning message

Wrong-Site Surgery Event- Situations in which one doctor scheduled to perform an operation is substituted by another occur commonly in hospitals. However, whether a doctor is an attending physician or not, it is possible that he/she might memorize the wrong operation site; this is particularly true of substitute doctors who do not know the patient well. Based on the scenario of wrong-site surgery, the system will automatically display the Ms. Wang's information including a chart on which the surgery site was marked while she was brought to OR-5 (See Fig.14). Therefore, the doctor was encouraged to review the displayed information and did not have to look for the paper-based chart.

Wrong-Patient Event-In the abovementioned situation, the patient —Shiau-ching Wang" probably underwent a wrong operation. However, with the RFID-based OR system, the nurse could be forced to reconfirm by using the admission application system.

When the nurse assigned the RFID wristband to the wrong patient —Shia-chin Wang," the system asked for the patient's ID card number. After the system compared the patient's ID card number with the OR schedule, a warning message would displayed (See Fig.18).

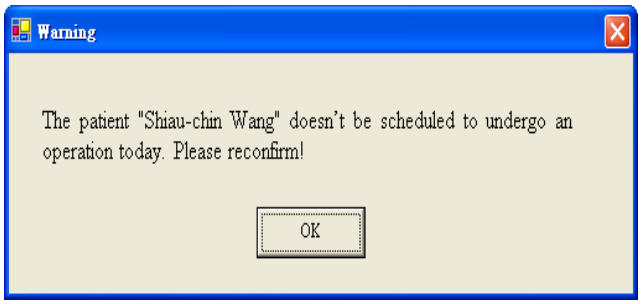


Fig.18 Warning message II

If the nurse incorrectly entered the ID card number of the correct patient —Shia-ching Wang," the system would display the correct patient's information including the patient's picture (See Fig.19). On seeing this displayed picture, the nurse would realize that a mismatch exists in the results and would therefore take corrective action.

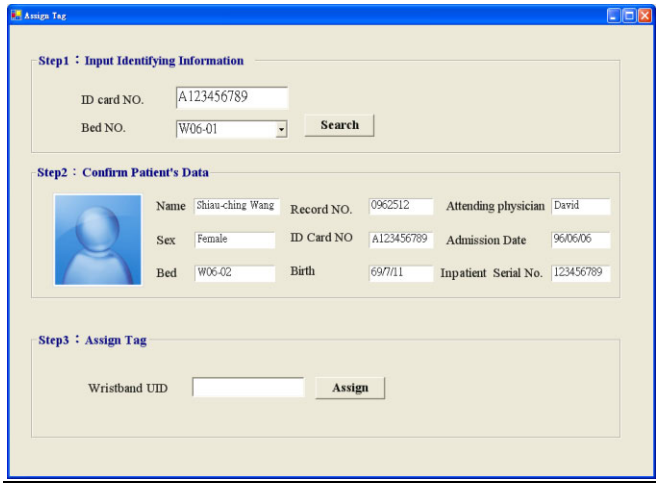


Fig.19 Reconfirmation by patient's picture

VI. CONCLUSION

Patient safety is the most important and uncompromised issue for medical institutions. In this research, we analyzed the existing OR process based on the as-is model and developed the RFID-based OR process. This RFID-based process can improve surgical patient safety and make medical staff efficient. From the surgical patient safety point of views, the RFID-based OR system (1) correctly identifies surgical patients, (2) automatically compares the OR which patients enter with the OR schedule, and (3) actively provides patients' information to ensure that patients get correct procedures. The proposed system decreases the probability of medical errors such as wrong patients, wrong locations, wrong medical staff and wrong procedures. From the medical staff point of views, the system replaces some time-wasted manual input processes. Therefore, it will improve operational efficiency in the OR and consequently help medical professionals better manage patient care.

VII. References

1) Martin A. Makary, Arnab Mukherjee, J. Bryan Sexton, Dora Syin, Emmanuelle Goodrich, Emily Hartmann, Lisa Rowen, Drew C. Behrens, Michael Marohn and Peter J. Pronovost(2007). Operating Room Briefings and Wrong-Site Surgery. Journal of the American College of Surgeons, 204(2), 236-243.
2) Mendelsohn D BSc MSc and Bernstein M MD MHSc(2009). Patient Safety in Surgery, Israeli Journal of Emergency Medicine, 9(2).
3) Deborah Carstens, Pauline Patterson, Rosemary Laird and Paula Preston(2009). Task analysis of healthcare delivery: A case study. Journal of Engineering and Technology Management, 26(1/2), 15-27.
4) Kohn, L. T., Corrigan, J. M., and Donaldson, M. S., editors(2000). To Err is Human. Building a Safer Health System. National Academy Press, Washington, DC.
5) Institute of Medicine, Committee on Quality of Health Care in America(2001). Crossing the Quality Chasm: A New Health Care System for the 21st Century, Washington, DC: National Academy Press.
6) The Joint Commission: The Joint Commission Home Page from http://www.jointcommission.org/
7) Gawande, A. A., Thomas E. J., Zinner M. J., and Brennan T. A.(1999). The incidence and nature of surgical adverse events in Colorado and Utah in 1992. Surgery, 126(1), 66-75.
8) Seiden, S. C., Barach, P.(2006), Wrong-Side/Wrong-Site, Wrong-Procedure, and Wrong-Patient Adverse Events, Archives of Surgery. 141(9), 931-939.
9) Bates, D. W.(2004). Using information technology to improve surgical safety. British journal of surgery, 91(8), 939-940.
10) Robin Riley and Elizabeth Manias(2009). Gatekeeping practices of nurses in operating rooms. Social Science & Medicine, 69(2), 215-222.

Diagnosis Of Heart Disease Using Datamining Algorithm

GJCST Classification
J.3

Asha Rajkumar[1], Mrs. G.Sophia Reena[2]

Abstract- **The diagnosis of heart disease is a significant and tedious task in medicine. The healthcare industry gathers enormous amounts of heart disease data that regrettably, are not "mined" to determine concealed information for effective decision making by healthcare practitioners. The term Heart disease encompasses the diverse diseases that affect the heart. Cardiomyopathy and Cardiovascular disease are some categories of heart diseases. The reduction of blood and oxygen supply to the heart leads to heart disease. In this paper the data classification is based on supervised machine learning algorithms which result in accuracy, time taken to build the algorithm. Tanagra tool is used to classify the data and the data is evaluated using 10-fold cross validation and the results are compared.**

Keywords: Naive Bayes, k-nn, Decision List, Tanagra tool

I. INTRODUCTION

The term heart disease applies to a number of illnesses that affect the circulatory system, which consists of heart and blood vessels. It is intended to deal only with the condition commonly called "Heart Attack" and the factors, which lead to such condition. Cardiomyopathy and Cardiovascular disease are some categories of heart diseases.The term ―cardiovascular disease" includes a wide range of conditions that affect the heart and the blood vessels and the manner in which blood is pumped and circulated through the body. Cardiovascular disease (CVD) results in severe illness, disability, and death. Narrowing of the coronary arteries results in the reduction of blood and oxygen supply to the heart and leads to the Coronary heart disease (CHD). Myocardial infarctions, generally known as a heart attacks, and angina pectoris, or chest pain are encompassed in the CHD. A sudden blockage of a coronary artery, generally due to a blood clot results in a heart attack. Chest pains arise when the blood received by the heart muscles is inadequate. High blood pressure, coronary artery disease, valvular heart disease, stroke, or rheumatic fever/rheumatic heart disease are the various forms of cardiovascular disease.

1) *Early Signs Of Heart Disease*

⇒ Dizzy spell or fainting fits.
⇒ Discomfort following meals, especially if long continued.
⇒ Shortness of breath, after slight exertion.
⇒ Fatigue without otherwise explained origin.
⇒ Pain or tightness in the chest a common sign of coronary insufficiency is usually constrictive in nature and is located behind the chest bone with
⇒ radiation into the arms or a sense of numbness or a severe pain in the centre of the chest.
⇒ Palpitation

II. DIAGONSIS OF HEART DISEASE

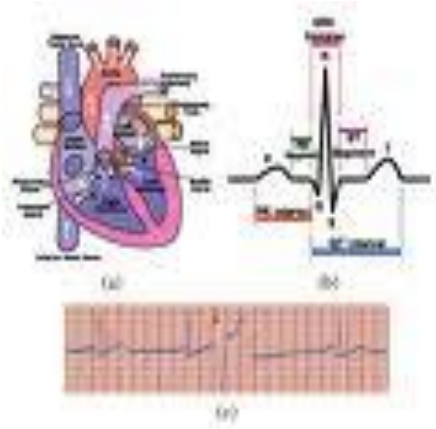


Fig.1 Diagnosis of heart disease

The initial diagnosis of a heart attack is made by a combination of clinical symptoms and characteristic electrocardiogram (ECG) changes. An ECG is a recording of the electrical activity of the heart. Confirmation of a heart attack can only be made hours later through detection of elevated **creatinine phosphokinase** (CPK) in the blood. CPK is a muscle protein enzyme which is released into the blood circulation by dying heart muscles when their surrounding dissolves.

1) *Risk factors for a heart attack are*
⇒ high blood pressure
⇒ diabetes
⇒ smoking
⇒ high cholesterol
⇒ family history of heart attacks at ages younger than 60 years, one or more previous heart attacks, male gender
⇒ obesity
⇒ Postmenopausal women are at higher risk than premenopausal women. This is thought to be due to loss of the protective effects of the hormone estrogen at menopause. It was previously treated by hormone supplements (hormone replacement therapy, or HRT).

About[1]-M.phil Scholar, P.S.G.R. Krishnammal College for Women, Coimbatore.
About[2]-HOD BCA Dept., P.S.G.R. Krishnammal College for Women, Coimbatore.Email id: ashar088@gmail.com

However, research findings have changed our thinking on HRT; long-term HRT is no longer recommended for most women.

⇒ Use of cocaine and similar stimulants.

III. EXTRACTION OF HEART DISEASE DATAWAREHOUSE

The heart disease data warehouse contains the screening the data of heart patients. Initially, the data warehouse is pre-processed to make the mining process more efficient. In this paper Tanagra tool is used to compare the performance accuracy of data mining algorithms for diagnosis of heart disease dataset. The pre-processed data warehouse is then classified using Tanagra tool. The feature selection in the tool describes the attribute status of the data present in the heart disease. Using supervised machine learning algorithm such as Naive Bayes, k-nn and Decision list and the result are compared.Tanagra is a collection of machine learning algorithms for data mining tasks. The algorithms can be applied directly to a dataset. Tanagra contains tools for data classification, statistics, clustering, supervised learning, meta-supervised learning and visualization. It is also well suited for developing new machine learning schemes. This paper concentrates on functional algorithms like Naive Bayes, k-nn, and Decision list.

IV. TANAGRA

Tanagra is a data mining suite build around graphical user interface. Tanagra is particularly strong in statistics, offering a wide range of uni- and multivariate parametric and nonparametric tests. Equally impressive is its list of feature selection techniques. Together with a compilation of standard machine learning techniques, it also includes correspondence analysis, principal component analysis, and the partial least squares methods. Tanagra is more powerful, it contains some supervised learning but also other paradigms such as clustering, supervised learning, meta supervised learning, feature selection, data visualization supervised learning assessment, statistics, feature selection and construction algorithms.The main purpose of Tanagra project is to give researchers and students an easy-to-use data mining software, conforming to the present norms of the software development in this domain , and allowing to analyze either real or synthetic data. Tanagra can be considered as a pedagogical tool for learning programming techniques. Tanagra is a wide set of data sources, direct access to data warehouses and databases, data cleansing, interactive utilization.

1) *Classification*

The basic classification is based on supervised algorithms. Algorithms are applicable for the input data. Classification is done to know the exactly how data is being classified. Meta-supervised learning is also supported which shows the list of machine learning algorithms. Tanagra includes support for arcing, boosting, and bagging classifiers. These algorithms in general operate on a classification algorithm and run it multiple times manipulating algorithm parameters or input data weight to increase the accuracy of the classifier. Two learning performance evaluators are included with Tanagra. The first simply splits a dataset into training and test data, while the second performs cross-validation using folds. Evaluation is usually described by accuracy, error, precision and recall rates. A standard confusion matrix is also displayed, for quick inspection of how well a classifier works.

2) *Manifold machine learning algorithm*

The main motivation for different supervised machine learning algorithms is accuracy improvement. Different algorithms use different rule for generalizing different representations of the knowledge. Therefore, they tend to error on different parts of the instance space. The combined use of different algorithms could lead to the correction of the individual uncorrelated errors. as a result the error rate and time taken to develop the algorithm is compared with different algorithm.

3) *Algorithm selection*

Algorithm is selected by evaluating each supervised machine learning algorithms by using supervised learning assessment (10-fold cross-validation) on the training set and selects the best one for application on the test set. Although this method is simple, it has been found to be highly effective and comparable to other methods. Several methods are proposed for machine learning domain. The overall cross validation performance of each algorithm is evaluated.

The selection of algorithms is based on their performance, but not around the test dataset itself, and also comprising the predictions of the classification models on the test instance. Training data are produced by recording the predictions of each algorithm, using the full training data both for training and for testing. Performance is determined by running 10-fold cross-validations and averaging the evaluations for each training dataset. Several approaches have been proposed for the characterization of learning domain. the performance of each algorithm on the data attribute is recorded. The algorithms are ranked according to their performance of the error rate.

4) *Manuscript details*

This paper deals with Naive Bayes, K-nn, Decision List algorithm . Experimental setup is discussed using 700 data and the results are compared . The performance analysis is

done among these algorithms based on the accuracy and time taken to build the model.

V. ALGORITHM USED

A Bayes classifier is a simple probabilistic classifier based on applying Bayes theorem with strong (naive) independence assumptions. In simple terms, a naive Bayes classifier assumes that the presence (or absence) of a particular feature of a class is unrelated to the presence (or absence) of any other feature. Depending on the precise nature of the probability model, naive Bayes classifiers can be trained very efficiently in a supervised learning setting. In many practical applications, parameter estimation for naive Bayes models uses the method of maximum likelihood. The advantage of using naive bayes is that one can work with the naive Bayes model without using any Bayesian methods.Naive Bayes classifiers have works well in many complex real-world situations. An advantage of the naive Bayes classifier is that it requires a small amount of training data to estimate the parameters (means and variances of the variables) necessary for classification. Because independent variables are assumed, only the variances of the variables for each class need to be determined and not the entire covariance matrix.A decision list has only two possible outputs, *yes* or *no* (or alternately 1 or 0) on any input. a decision list is a question in some formal system with a yes-or-no answer, depending on the values of some input parameters . A method for solving a decision problem given in the form of an algorithm is called a decision procedure for that problem. The field of computational complexity categorizes decidable decision problem. Research in computability theory has typically focused on decision problems These inputs can be natural numbers, also other values of some other kind, such as strings of a formal language. Using some encoding, such as Godel numberings, the strings can be encoded as natural numbers. Thus, a decision problem informally phrased in terms of a formal language is also equivalent to a set of natural numbers. To keep the formal definition simple, it is phrased in terms of subsets of the natural numbers. Formally, a decision problem is a subset of the natural numbers. The corresponding informal problem is that of deciding whether a given number is in the set. Decision problems can be ordered according to many-one reducibility and related feasible reductions such as Polynomial-time reductions. Every decision problem can be converted into the function problem of computing the characteristic function of the set associated to the decision problem. If this function is computable then the associated decision problem is decidable. However, this reduction is more liberal than the standard reduction used in computational complexity (sometimes called polynomial-time many-one reduction). The *k*-nearest neighbor‘s algorithm (*k*-NN) is a method for classifying objects based on closest training data in the feature space. k-NN is a type of instance-based learning. The function is only approximated locally and all computation is deferred until classification. The *k*-nearest neighbor algorithm is amongst the simplest of all machine learning algorithms. The same method can be used for regression, by simply assigning the property value for the object to be the average of the values of its *k* nearest neighbors. It can be useful to weight the contributions of the neighbors, so that the nearer neighbors contribute more to the average than the more distant ones The neighbors are taken from a set of objects for which the correct classification (or, in the case of regression, the value of the property) is known. This can be thought of as the training set for the algorithm, though no explicit training step is required. The *k*-nearest neighbor algorithm is sensitive to the local structure of the data Nearest neighbor rules in effect compute the decision boundary in an implicit manner. It is also possible to compute the decision boundary itself explicitly, and to do so in an efficient manner so that the computational complexity is a function of the boundary complexity. The best choice of *k* depends upon the data; generally, larger values of *k* reduce the effect of noise on the classification, but make boundaries between classes less distinct. A good *k* can be selected by various heuristic techniques, for example, cross-validation. The special case where the class is predicted to be the class of the closest training sample (i.e. when $k = 1$) is called the nearest neighbor algorithm.The accuracy of the *k*-NN algorithm can be severely degraded by the presence of noisy or irrelevant features, or if the feature scales are not consistent with their importance. Much research effort has been put into selecting or scaling features to improve classification. In binary (two class) classification problems, it is helpful to choose *k* to be an odd number as this avoids tied votes.Using an appropriate nearest neighbor search algorithm makes *k*-NN computationally tractable even for large data sets. The nearest neighbor algorithm has some strong consistency results. As the amount of data approaches infinity, the algorithm is guaranteed to yield an error rate no worse than twice the Bayes error rate (the minimum achievable error rate given the distribution of the data. *k*-nearest neighbor is guaranteed to approach the Bayes error rate, for some value of *k* (where *k* increases as a function of the number of data points). Various improvements to *k*-nearest neighbor methods are possible by using proximity graphs.The training data set consists of 3000 instances with 14 different attributes. The instances in the dataset are representing the results of different types of testing to predict the accuracy of heart disease. The performance of the classifiers is evaluated and their results are analyzed.In general, tenfold cross validation has been proved to be statistically good enough in evaluating the performance of the classifier. The 10-fold cross validation was performed to predict the accuracy of heart disease. The purpose of running multiple cross-validations is to obtain more reliable estimates of the risk measures. The basic crisis is to predict the accuracy of heart disease that can be stated as follows: particular dataset of heart disease with its appropriate attributes. The main aim is to get a accuracy by classifying algorithms.

VI. EXPERIMENTAL SETUP

The data mining method used to build the model is classification. The data analysis is processed using Tanagra data mining tool for exploratory data analysis, machine

learning and statistical learning algorithms. The training data set consists of 3000 instances with 14 different attributes. The instances in the dataset are representing the results of different types of testing to predict the accuracy of heart disease. The performance of the classifiers is evaluated and their results are analysed. The results of comparison are based on 10 ten-fold cross-validations. According to the attributes the dataset is divided into two parts that is 70% of the data are used for training and 30% are used for testing.

1) *Description of dataset*

The dataset contains the following attributes:

1) id: patient identification number
2) age: age in year,
3) sex: sex (1 = male; 0 = female),
4) painloc: chest pain location (1 = substernal; 0 = otherwise),
5) painexer (1 = provoked by exertion; 0 = otherwise),
6) relrest (1 = relieved after rest; 0 = otherwise),
7) cp: chest pain type
   - ⇒ Value 1: typical angina
   - ⇒ Value 2: atypical angina
   - ⇒ Value 3: non-anginal pain
   - ⇒ Value 4: asymptomatic
8) trestbps: resting blood pressure
9) chol: serum cholestoral
10) famhist: family history of coronary artery disease (1 = yes; 0 = no)
11) restecg: resting electrocardiographic results
    - ⇒ Value 0: normal
    - ⇒ Value 1: having ST-T wave abnormality (T wave inversions and/or ST
    - ⇒ elevation or depression of > 0.05 mV)
    - ⇒ Value 2: showing probable or definite left ventricular hypertrophy
    - ⇒ by Estes' criteria
12) ekgmo (month of exercise ECG reading)
13) thaldur: duration of exercise test in minutes
14) thalach: maximum heart rate achieved
15) thalrest: resting heart rate
16) num: diagnosis of heart disease (angiographic disease status)
    - ⇒ Value 0: < 50% diameter narrowing
    - ⇒ Value 1: > 50% diameter narrowing
17) (in any major vessel: attributes 59 through 68 are vessels)

2) *Learning Algorithms*

This paper consists of three different supervised machine algorithms learning algorithms derived from the Tanagra data mining tool. Which include:

- ⇒ naive bayes,
- ⇒ k-nn,
- ⇒ decision list

The above algorithms were used to predict the accuracy of heart disease.

3) *Performance study of Algorithms*

The table I consists of secondary values of different classifications. According to these values the accuracy is calculated and analyzed. It has 14 attributes for classification. Each one has a distinct value. Performance can be determined based on the evaluation time of calculation and the error rates. Comparison is made among these classification algorithms out of which the naive bayes algorithm is considered as the better performance algorithm. Because it takes only some time to calculate the accuracy than other algorithms .

Table I. Performance Study Of Algorithm

Algorithm Used	Accuracy	Time Taken
Naive Bayes	**52.33%**	**609ms**
Decision List	52%	**719ms**
KNN	**45.67%**	**1000ms**

Naive bayes algorithm results in lower error ratios than decision list and knn algorithm that we have experimented with a training dataset.

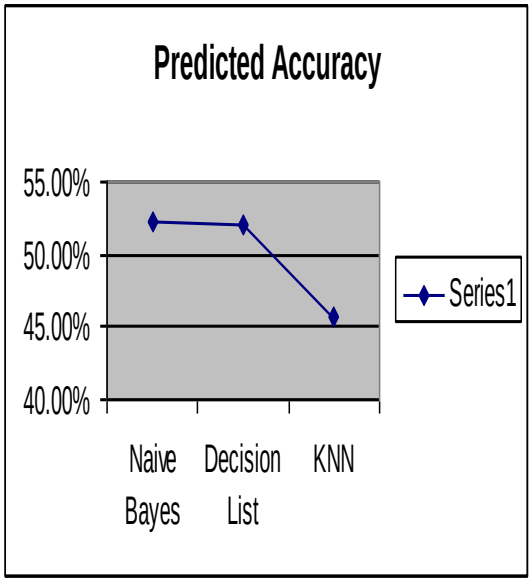


Fig.2 Predicted Accuracy

Fig.2 represents the resultant values of above classified datasets using data mining supervised machine learning algorithms and it shows the error ratio of three different algorithms, which was compared. It may vary according to it attributes of heart disease dataset. It is logical from the

chart that naive bayes algorithm shows the superior performance compared to other algorithms.

Table Ii. Performance Study Of Accuracy

Algorithm used	Values	Recall	1-precision
Naïve Bayes	**Left vent hypertrophy**	**0.4828**	**0.4853**
	normal	**0.5705**	**0.4753**
	St-t-abnormal	**0.0000**	**1.0000**
Decision List	Left vent hypertrophy	0.4897	0.4855
	normal	0.5705	0.4688
	St-t-abnormal	0.0000	1.0000
KNN	Left vent hypertrophy	0.4552	0.5479
	normal	0.4765	0.5390
	St-t-abnormal	0.0000	1.0000

The table II consists of different values. According to these values the accuracy is calculated using three main attribute namely left ventricle hypothesis, normal and stress abnormal. Performance can be determined based on the comparison of the accuracy. Comparison is made among these classification algorithms out of which the Naive Bayes algorithm is considered as the better performance algorithm.

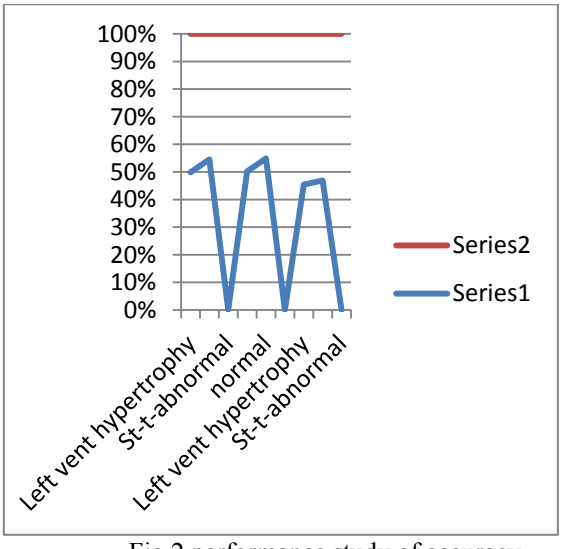


Fig.2 performance study of accuracy

Fig.2 represents the resultant values of above classified datasets using data mining supervised machine learning algorithms and it shows the accuracy of three different algorithms, which was compared. It may vary according to it attributes of heart disease dataset. It is logical from the chart that Naive Bayes algorithm shows the superior performance compared to other algorithms.

VII. Conclusion

Data mining in health care management is unlike the other fields owing to the fact that the data present are heterogeneous and that certain ethical, legal, and social constraints apply to private medical information. Health care related data are voluminous in nature and they arrive from diverse sources all of them not entirely appropriate in structure or quality. These days, the exploitation of knowledge and experience of numerous specialists and clinical screening data of patients gathered in a database during the diagnosis procedure, has been widely recognized. This paper deals with the results in the field of data classification obtained with Naive Bayes algorithm, Decision list algorithm and k-nn algorithm, and on the whole performance made known Naive Bayes Algorithm when tested on heart disease datasets. Naive Bayes algorithm is the best compact time for processing dataset and shows better performance in accuracy prediction. The time taken to run the data for result is fast when compared to other algorithms. It shows the enhanced performance according to its attribute. Attributes are fully classified by this algorithm and it gives 52.33% of accurate result. Based on the experimental results the classification accuracy is found to be better using Naive Bayes algorithm compare to other algorithms. From the above results Naive Bayes algorithm plays a key role in shaping improved classification accuracy of a dataset

VIII. Reference

1) Chen, J., Greiner, R.: Comparing Bayesian Network Classifiers. In Proc. of UAI-99, pp.101–108 ,1999.
2) Quinlan, J.: C4.5: Programs for Machine Learning. Morgan Kaufmann, San Mateo 1993.
3) "Hospitalization for Heart Attack, Stroke, or Congestive Heart Failure among Persons with Diabetes", Special report:
2001 – 2003, New Mexico.
4) M. Chen, J. Han, and P. S. Yu. Data Mining: An Overview from Database Perspective. IEEE Trans. Knowl. Dat. Eng., vol: 8, no:6, pp: 866-883, 1996.
5) R. Agrawal and R. Srikant, "Fast algorithms for mining association rules in large databases", In Proceedings of the 20th International Conference on Very Large Data Bases, Santiago, Chile, August 29-September 1994.
6) Niti Guru, Anil Dahiya, Navin Rajpal, "Decision Support System for Heart Disease Diagnosis Using Neural Network", Delhi Business Review , Vol. 8, No. 1 (January - June 2007).
7) Carlos Ordonez, "Improving Heart Disease Prediction Using
Constrained Association Rules," Seminar Presentation atUniversity of Tokyo, 2004.
8) Franck Le Duff , Cristian Munteanb, Marc Cuggiaa, Philippe Mabob, "Predicting Survival Causes After Out of Hospital Cardiac Arrest using

Data Mining Method", Studies in health technology and informatics ,107(Pt 2):1256-9, 2004.

9) Boleslaw Szymanski, Long Han, Mark Embrechts, Alexander Ross, Karsten Sternickel, Lijuan Zhu, "Using Efficient Supanova Kernel For Heart Disease Diagnosis", proc. ANNIE 06, intelligent engineering systems through artificial neural networks, vol. 16, pp:305-310, 2006.
10) Agrawal, R., Imielinski, T. and Swami, A, _Mining association rules between sets of items in large databases', Proc. ACM-SIGMOD Int. Conf. Management of Data (SIGMOD'93),Washington, DC, July, pp.45–51, 1993.
11) "Heart Disease" from http://chineseschool. netfirms.com/heart-disease-causes.html
12) "Heart disease" from http://en.wikipedia.org/wiki/Heart_disease.
13) Kiyong Noh, Heon Gyu Lee, Ho-Sun Shon, Bum Ju Lee,and Keun Ho Ryu, "Associative Classification Approach for Diagnosing Cardiovascular Disease", springer,Vol:345 , pp 721-727, 2006.
14) Sellappan Palaniappan, Rafiah Awang, "Intelligent Heart Disease Prediction System Using Data Mining Techniques", IJCSNS International Journal of Computer Science and Network Security, Vol.8 No.8, August 2008
15) Andreeva P., M. Dimitrova and A. Gegov, Information Representation in Cardiological Knowledge Based System,SAER'06, pp: 23-25 Sept, 2006.

Information Security Risk Assessment for Banking Sector-A Case study of Pakistani Banks

Usman Munir[1], Irfan Manarvi[2]

GJCST Classification
E.5 J.1 H.2.7

***Abstract*- The ever increasing trend of Information Technology (IT) in organizations has given them new horizon in international market. Organizations now totally depend on IT for better and effective communication and daily operational tasks. Advancements in IT have exposed organization to information security threats also. Several methods and standards for assessment of information security in an organization are available today. Problems with these methods and standards are that they neither provide quantitative analysis of information security nor access potential loses information malfunctioning could create. This paper highlight the necessity of information security tool which could provide quantitative risk assessment along with the classification of risk management controls like management, operational and technical controls in an organizations. It is not possible for organizations to establish information security effectively without knowing the loopholes in their controls. Empirical data for this research was collected from the 5 major banks of Pakistan through two different questionnaires. It is observed that mostly banks have implemented the technical and operational control properly, but the real crux, the information security culture in organization is still a missing link in information security management.**

Keywords– Quantitative information security assessment, information security controls, information security, information security management system, risk, risk management.

I. INTRODUCTION

Information is considered as an asset like other important business assets and Information Security (IS) is a way of protecting information from a wide range of threats in order to ensure business continuity, minimize risk, and maximize return on investments and business opportunities[1 and 2]. Over the years the usage of Information Technology (IT) has increased massively in organizations and in society and to cater the ever increasing requirement of information flow, information systems has become complex and multifaceted [4]. IT has made electronic communication and internet necessary in all organizations. This necessity has brought efficiency and threats of hacking and intrusion with it [5]. With all these advancements in the field of IT, dependency of organizational business fusnctionality on it has increased the requirement of securing organizational information from threats [3][4][6][8]. Information security is somewhat a hard task to achieve. One of the prime reasons is that not much data related to information security management and threats to organizations‘

About[1]- Department of Engineering Management Department, CASE, Islamabad, Pakistan
About[2]- Department of Mechanical Engineering, HITEC University, Taxila, Pakistan

information is available due to confidentiality [12]. Second, costs associated to information security restrict organizations from implementing information security management systems in organizations [13]. Third, information security is not just a technical issue, it is more a managerial issue, therefore it is also required to train employees about the information security without which attaining information security is impossible [6]. It is proposed implementing information security management policy in such a way that it calculates the assets values first and then predicts the losses associated to it [10].But so far available quantitative risk assessment methods and tools are either expensive or little information about their usability and performance are available [9]. Although operational risk management techniques are functional in many originations but it is not possible to handle information risks with the perspective of operational risk management. It is therefore advised combining information security management with operational risk management for a better economical solution [7]. More dependency of world on information technology systems and processes made management of information technology risk a practical necessity [14]. Organizations adopt information security product, services, processes and tools which range from complex mathematical algorithms to the expert risk management resources. Organizations are not sure about the optimal security quality and required a cost effective information security methods which can provide them optimal security with minimum cost.Knowledge sharing and collaboration of intra-organizational cross functional teams for risk management is required for proper risk management strategies [15]. Management vision towards information security risk and involving internal stakeholder in this task is the need of the time. A more pragmatic reason is that the development of information security methods within organisations is rather an ad hoc process than a systematic one. This process generates new knowledge about information security risk management by constituting valuable organizational intelligence. Therefore, it is very important to have a systematic process, which ensures that the acquired knowledge will be elicited, shared, and managed appropriately [12].Codification and personalization are two strategies for information security. Codification is the people to document strategy to ensure intranets and databases are loaded with best practices, case studies and guidance for people in their day-to-day work. Personalization is the people-to-people strategy to link people and grow network and information security culture [16]. These two strategies established the reusability of implemented processes in organization and information

sharing background in organization. A creditable and effective method for accessing current state of information security in organizations is desirable. Good information about the current system required for good decision. Assessment of current method would help in future improvement in the system traded by its implementation cost [17]. Information security of an enterprise was defined in term of tree structure [18]. Prioritized the structure on enterprise information security basis [19], to clarify the assessment scope and minimize the assessment cost. Finally, the credibility of the assessment results is addressed with a statistical approach combined with ideas from historical research and witness interrogation psychology [20].

II. Information Security Risk Management Tool: COBRA™

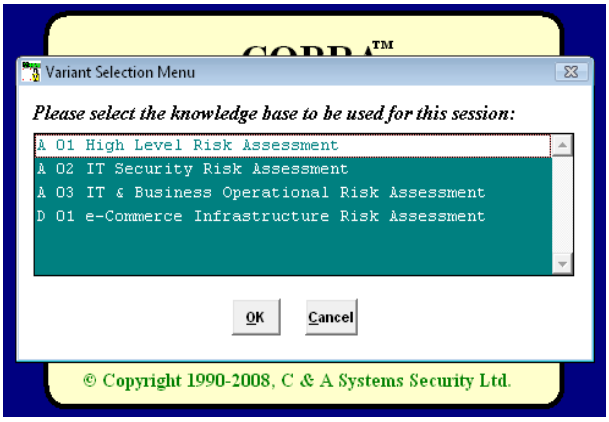
A number of standards are available on information security management system like ISO 17799, ISO 27001, the Control Objectives for Information and related technology (Cobit), and National Institute of Standards and Technology (NIST). These standards describe the requirement of information security in the organizations. But so far experts for information security implementation are requires. Whose services are expensive and rely only on their judgment for information security risk management process. Available tools are also expensive and generic. COBRA™ risk consultant is software which provides the following risk assessment:

- Compliance with the ISO information security standards BS7799 and 17799.
- Support implementing Information Security Management System in organization and also assess risk associated with organization information.
- Quantitative risk assessment of information security threats.
- Supported with a built-in knowledge base which acts as a database for evaluating information security risks.
- Perform risk assessment and also suggest its mitigation approaches.
- Assess Business Continuity Plan of an organization against its profile.
- Provide comprehensive reports about the information security risk.

Four knowledge base available to perform risk assessment are as follows and shown in Fig 1:

1. High level Risk Assessment
2. IT Security Risk Assessment
3. IT & Business operational Risk Assessment
4. E-commerce Infrastructure Risk Assessment

Fig 1: COBRA™ front-end



For the risk assessment a questionnaire have to be filled by a respondent which then generate a comprehensive report of its organizational risk.

III. Research Methodology

This study was conducted to analyze quantitative risks associated with information of major banks operating in Pakistan and to study security control, which are implemented for protecting their critical information. Two separate questionnaires were designed for the analysis of information security and its controls. First questionnaire was developed by using built-in knowledge base from High Level Risk Assessment section of COBRA™ software. After reviewing the COBRA™ software analysis another questionnaire was designed to evaluate the management control in these banks.High Level Risk Assessment questionnaire was filled by the information security auditors and by higher management of these banks to get the response about their implemented security policies and its impact on their information security management. Second questionnaire was filled by the five persons of various management level of each bank to check the management vision toward the information security and awareness of information security in these banks. To perform information security analysis, threats to information confidentiality, integrity, and availability were checked against the information security controls like management, technical, and operational controls.

IV. Quantitative Analysis of COBRA™ Questionnaire

COBRATM asked questions to evaluate potential threats and security policies in an organization. To determine the high level risk related to the information of the organization, questions were classified into the four categories:

a. Availability
b. Business Impact Analysis
c. Confidentiality
d. Integrity

Few important questions of COBRATM software along with their assigned points are discussed bellow to observe its functionality.

1) *Availability Questionnaire*

In availability section the questions will the highest value are discussed in table 1 below:

Table 1: Availability questionnaire

No	Availability	Answer	Scr
1	Is there a formal and workable Business Redemptions Plan in place?	Yes	0
		No	50
2	How confident are you that the plan is adequate to ensure a controlled recovery and continuance of business within the time frames specified as significant/critical:	100 % Confident	0
		Fairly Confident	0.5
		Comfortable	1
		Concerned	20
		Not Confident	50
3	When the Business Continuity Plan was last tested?	Within the 12 months	0
		1-2 years	0.5
		2-3 years	1
		4-5 years	20
		more than 5 years ago	50
4	Are the contingency arrangements for all key components reasonable and appropriate?	Yes	0
		No	50
5	How confident are you that the contingency arrangements and Business Continuity Plan would enable continuance and eventual recovery from the loss of a key building (due perhaps to serious fire, flooding, explosion, etc) without serious or critical impact on the business?	100 % confident	0
		Fairly Confident	0.5
		Comfortable	1
		Not Really Confident	20
		Concerned	50
6	How confident are you that the contingency arrangements and Business Continuity Plan would enable continuance and eventual recovery from the loss of key personnel (due perhaps to serious accident, industrial action, etc) without serious or critical impact on the business?	100 % confident	0
		Fairly Confident	0.5
		Comfortable	1
		Not Really Confident	20
		Concerned	50
7	Ignoring the recovery element of the Business Continuity Plan, to which of the following (if any), is the exposure level significant?	Fire/Flooding/ Explosion	20
		Hardware/ Equipment Malfunctin	20
		Hardware/ Equipment/ Media/Other	20
		Power Failure	20
		Software Error	20
		Infection By Computer Virus	20
		Intro Of Malicious Coding	20
8	Ignoring the recovery element of the Business Continuity Plan, to which of the following (if any), is the exposure level significant?	Hacking/Electronic Sabotage	20
		Loss Of 3rd Party Service	20
		Loss of Comm/ Network Service	20
		Operator Error /Sabotage	20
		Industrial Action by Key Staff	20
		Other Threat	20
9	Are specific back-up and recovery measures in place to handle both loss of critical data and serious software error in a timely and appropriate fashion?	Yes	0
		No	20
10	Are physical access controls/practices for areas that may hold sensitive/confidential information appropriate?	Certainly Adequate	0
		Generally OK	0.5
		A cause for concern	20
		A major problem	50

The questionnaire pertaining to Availability discussed the business continuity plan, business redemption, and disaster recovery plan. In this section maximum points were given to business continuity plan because it ensures the continuity of critical business functions by providing methods and procedures for dealing with long outages and disasters. It is a broader approach in which continuity of a business during any disaster is ensured until that disaster is either curtailed or business operation returns to its normal circumstances. Physical access controls were also evaluated because appropriate physical controls are necessary to eliminate potential losses and risks associated to information assets, weak physical access controls could not prevent intruder

from causing any harm to information processing facility or information assets. Therefore, COBRA™ ask particular questions related to appropriate physical access control not only to evaluate these control for better analysis of security situation but also to create awareness in management for importance of these controls.

2) *Business Impact Questionnaire*

Business Impact considered as a functional analysis in which a team collects data through interviews and documentary resources. Than developing hierarchy of business functions and applies a classification scheme to indicate each individual function critical level. Question from the Business Impact Analysis are shown bellow in table 2.

Table 2: Business Impact questionnaire

No	Business Impact	Answer	Scr
11	What was the total revenue for this business function/service during the last financial year?	Less Than 10,000,000	0
		10,000,000 to 100,000,000	1
		100,000,000 to 500,000,000	5
		More than 500,000,000	20
12	What is the highest likely financial value throughout per day :	Less than 500,000	0
		500,000 to 5,000,000	1
		5,000,000 to 50,000,000	5
		More than 50,000,000	20
13	Which of the following types of function are directly performed :	Financial Accounting	5
		Trading/ Dealing	5
		Payroll	5
		Management info/ Support	6
		Research	5
		Manufacturing	5
		Infra-structure Support	5
		Retail	5
		Other	5
14	How many other systems or business units internal to this enterprise have a dependency upon this one?	Minor Dependency	1
		Significant Dependency	2
		Total Dependency	3
15	In the worst case scenario means no backup, how quickly could unavailability result in SIGNIFICANT impact in terms of current/future revenues and other direct financial losses?	2 hours	200
		24 hours	200
		7days	2
		1 month	1
		Never	0
16	In the worst case scenario, how quickly could unavailability have a SIGNIFICANT impact in terms of customer, shareholder, public or departmental confidence?	2 hours	999
		24 hours	200
		7days	20
		1 month	1
		Never	0
17	How quickly could unavailability have a SIGNIFICANT impact in terms of contractual, regulatory, or legal obligations?	2 hours	200
		24 hours	20
		7days	2
		1 month	1
		Never	0
18	If confidential/key information was disclosed to one or more competitors, what is the worst impact that could result:	None	999
		Moderate	200
		Significant	20
		Substantial	1
		Critical	0
19	If confidential/key information was disclosed, what could be the worst impact in terms of current/future revenues and other direct financial losses?	None	200
		Moderate	20
		Significant	2
		Substantial	1
		Critical	0
20	If confidential/key information was disclosed, what could the worst impact be in terms of customer, shareholder, public or departmental confidence?	None	999
		Moderate	200
		Significant	20
		Substantial	1
		Critical	0
21	If confidential/key information was disclosed, would there be any implications in terms of contractual, regulatory, or legal obligations?	None	0
		Moderate	1
		Significant	5
		Substantial	25
		Critical	150
22	If the data/information lost its integrity (through error, deliberate unauthorized alteration, fraud, etc), what could be the worst impact in terms of direct financial loss?	None	0
		Moderate	1
		Significant	5
		Substantial	25
		Critical	150
23	If the data/information lost its integrity, what could the worst impact be in terms of customer, shareholder, public or departmental confidence?	None	0
		Moderate	1
		Significant	5
		Substantial	25
		Critical	150

24	If the data/information lost its integrity, would there be any implications in terms of contractual, regulatory, or legal obligations?	None	0
		Moderate	1
		Significant	5
		Substantial	25
		Critical	150

In business impact analysis section the COBRA™ gave maximum points to questions which have direct and indirect impact on the stakeholders and customers. The confidence of customer and internal and external stakeholders on business products and business management process is essential for the success. All these questions are concerned on unavailability of critical information to business in term of customers and stakeholders confidence on the organization. COBRA™ also bifurcate impact of losing critical information with respect to time to establish minimum time when the organization would have maximum disadvantage of losing that information. The confidence of organization or management on its internal and external employees and stakeholders is considered as a key to establish proper information security checks.

3) *Confidentiality Questionnaire*

This section of software is about establishing confidentiality of information. Assurance that the information is not disclosed to any unauthorized individual, programs or processes. Organizations implement information confidentiality separately for business viability. Questions and points assigned by COBRA™ are discussed in table 3:

Table 3: Confidentiality questionnaire

No	Confidentiality	Answer	Scr
25	How confident are you that there is no serious threat of a third party having unauthorized sight of sensitive hardcopy output?	100 % confident	0
		Fairly Confident	0.5
		Comfortable	1
		Not Really Confident	20
		Concerned	50
26	Are physical access controls/practices for the building appropriate?	Certainly	0
		Okay	0.5
		Cause of Concern	20
		Major Problem	50
27	Are physical access controls/practices for areas that may hold sensitive/confidential information appropriate?	Certainly	0
		Okay	0.5
		Cause of Concern	20
		Major Problem	50
28	Are logical access controls sufficient to protect sensitive data/information from unauthorized EXTERNAL scrutiny?	No weakness	0
		Minor Weakness	0.5
		Not Sure	1
		Some Concerns	20
		Major Weakness	50
29	Are logical access controls appropriate and sufficient to protect sensitive data/information from unauthorized INTERNAL scrutiny?	No weakness	0
		Minor Weakness	0.5
		Not Sure	1
		Some Concerns	20
		Major Weakness	50
30	Are practices with respect to hardware, equipment and media adequate and appropriate?	Yes	0
		No	50
31	Is the security infra-structure and culture of the enterprise:	Good	0.5
		Reasonable	1
		Poor	50
32	Are there any other exposures evident?	No Major Exposure	0
		Some concerns	20
		Significant Exposure	20
33	Are there measures/plans in place to mitigate or manage any breach of confidentiality?	Yes	0
		Outlines only	0.5
		Ideas Only	2
		Nothing	50

COBRA™ gives emphasis on questions related to the physical and logical access controls because in the current scenarios when companies have threat over internet about the security breach and intrusion, lapses on these parts can harm the organization in term of market repute and profitability. Besides these other important questions are related to security structure and culture of an enterprise. Information security is based mostly on culture of an organization. If in an organizations every employee is aware about the information security requirement, policies, and have good understanding of their responsibilities towards information security, organization will have less threats of losing information than.

4) *Integrity Questionnaire*

Integrity of information means protecting data and information resource from being altered in an unauthorized fashion. The questions in the integrity section are assigned the following scores as in table 4:

Table 4: Integrity questionnaire

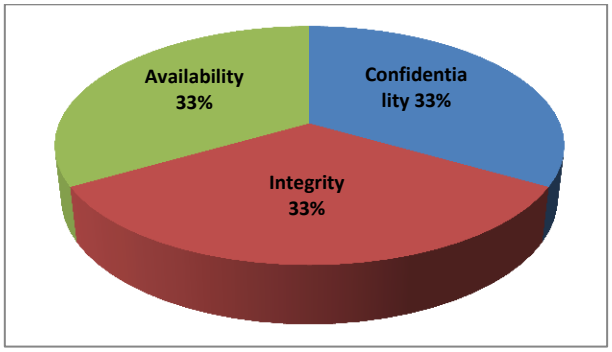
No	Integrity	Answer	Scr
34	How confident are you that there is no significant risk of serious ERROR being introduced during the input of important data/information?	100 % confident	0
		Fairly Confident	0.5
		Comfortable	1
		Not Really Confident	20
		Concerned	50
35	Consider the situation with respect to INTENTIONAL unauthorized manipulation of input data/information, by both internal and external parties. How confident are you that there is no significant risk of serious breach during the input of important data/information?	100 % confident	0
		Fairly Confident	0.5
		Comfortable	1
		Not Really Confident	20
		Concerned	50
36	How confident are you that there is no significant risk of serious error being introduced via program error or malfunction?	100 % confident	0
		Fairly Confident	0.5
		Comfortable	1
		Not Really Confident	20
		Concerned	50
37	Are the controls in place to prevent the unauthorized modification of program source code appropriate?	Certainly	0
		Okay	0.5
		Cause of Concern	20
		Major Problem	50
38	Are logical access controls sufficient to protect sensitive data/information from unauthorized EXTERNAL access?	No weakness	0
		Minor Weakness	0.5
		Not Sure	1
		Some Concerns	20
		Major Weakness	50
39	Are logical access controls appropriate and sufficient to protect sensitive data/information from unauthorized INTERNAL access?	No weakness	0
		Minor Weakness	0.5
		Not Sure	1
		Some Concerns	20
		Major Weakness	50
40	Are the controls over computer operations adequate and appropriate?	Certainly	0
		Okay	0.5
		Cause of Concern	20
		Major Problem	50
41	Is the security infra-structure and culture of the enterprise:	Excellent	0
		Good	0.5
		Reasonable	1
		Poor	50
42	Are there any other exposures evident?	No Major Exposure	0
		Some concerns	20
		Significant Exposure	50

In this section the COBRA™ asks management about their confidence on internal and external security checks which are implemented to save data from any unofficial changes. Organizations mostly have external and internal threats to their information. The internal threat can lead to more catastrophic impact than the external one. Therefore it is important for the organization to have a full confidence on the internal security controls and procedures on retrieving and adding data. If internal procedure and process of input and output of information has some loopholes then it is important to adjust and redefine these procedure and make it as firm as required for the integrity of information. As a result organizations assign different privileges to users so that only designated officials can alter or process information. These controls help organization in maintaining the security checks and also give sense of responsibility to the personals authorized for any changes.

5) *Bifurcation of Information Threats in COBRA™*

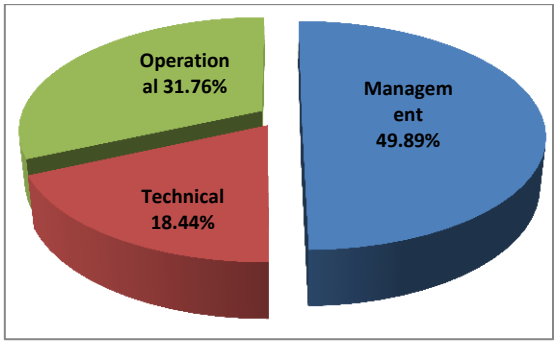
Besides the importance of some particular questions on others COBRA™ has given equal score to availability, integrity and confidentiality of information as shown in Fig 2 that all parts have given equal percentage of 33%.

Fig 2: Bifurcation of C.I.A in COBRA™



To evaluate organization information security controls, COBRA™ questions were plotted against three information security controls. By this, it has observed that COBRA™ constructed questions in such a format that the major distribution of these questions were related to management control with a percentage of 49.89% following the operational control with 31.76% and finally technical control with 18.44% as shown in Fig 3.

Fig 3: Bifurcation of security controls in COBRA™

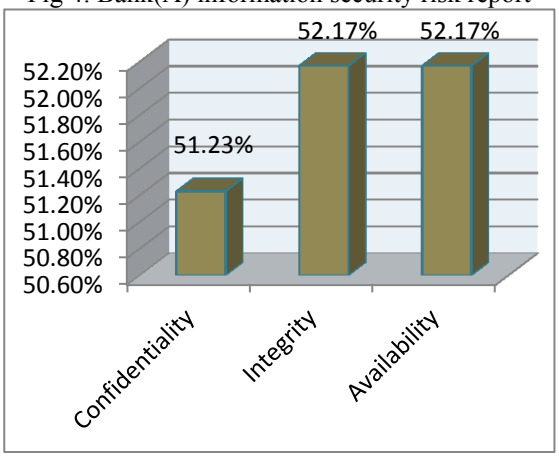


V. COBRA™ APPLICATION IN BANKING SECTOR

1) *High Level Risk Assessment of Bank (A)*

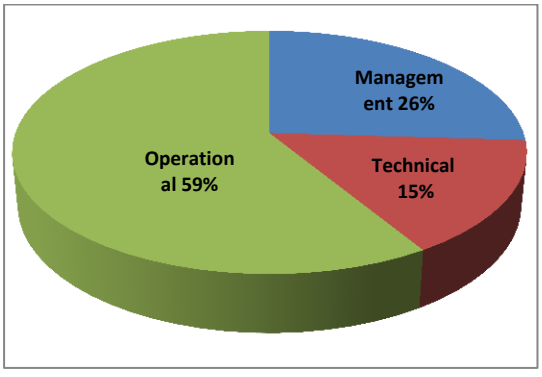
Bank (A) is one of the largest banks of Pakistan with presence in Hong Kong, UK, Nepal, Nigeria, Kenya and Kyrgyzstan and rep offices in Iran and China. Key areas of operations encompass product offerings and services in retail and consumer banking.The analysis done by the COBRA™ on threats to information availability, integrity and confidentiality is shown in Fig 4. According to it, threat to information confidentiality is 51.23% showed that the information could be intruded. Whereas threat to integrity of information is 52.37% means information could be altered or in some cases a completeness of information was questionable for organization. Threat to availability of information in this bank was at 52.17% too. This software suggested implementing security checks on data warehouses and physical data places.

Fig 4: Bank(A) information security risk report



The figure 5 shows the percentage of security controls in bank (A). It may be seen that the management control is lower being 26% against suggested 49.89%. The operational control is being at 59% where the suggested percentage at 31.76%. Technical control being lower than the suggested 18.44% was at 15%. This showed that the bank overall operational threats were catered properly whereas the implementation of information security policy was uncertain and lack of information security awareness in the management functions.

Fig 5: Bank(A) information security controls report



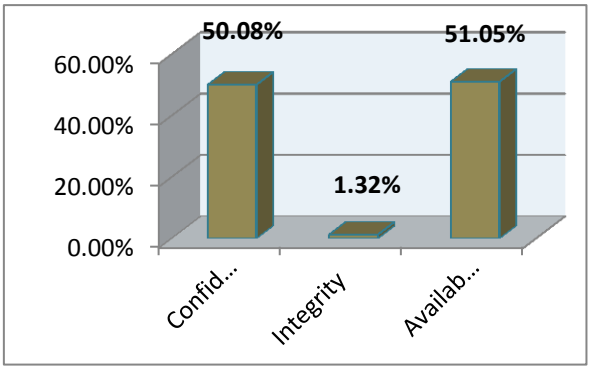
2) *High Level Risk Assessment of Bank (B)*

It is the fourth largest bank of Pakistan has a customer base of approximately 4 million, 1,026 branches, and over 300 ATMs. The bank (B) risk assessment through COBRA™ shown in Fig 6. The threats to confidentiality of information were at 50.08% showed an unauthorized access of data. The threats to integrity is maintained well which were at 1.32% only. Threats to availability of information were reported 51.05 % showed concern to information record keeping.

Precise review of report showed that information integrity related to the information accuracy was well maintained in the bank. But other information security threats like confidentiality and availability were quite high. These high levels of threats showed that the availability of critical information to unauthorized and unwanted individuals or to

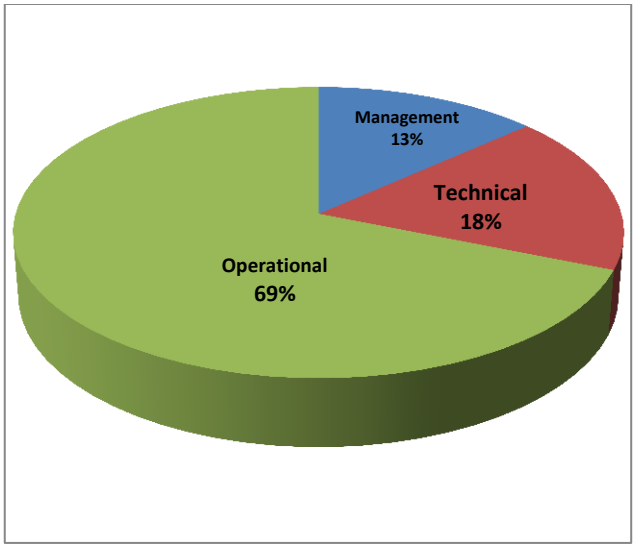
the third-part. This can harm the reputation of the bank and ultimately can affect its business.

Fig 6: Bank(B) information security risk report



The Fig 7 showed the percentage of security controls in bank (B). It may be seen that the management control was lower being 13% showed a high threats related to it. The technical control and operational controls were optimal being at 18% and at 69%. This showed an improper management control in this bank. It led to a risk of improper information security policy and its implementation, information leakage by employees, unsecure risk culture in organization, and unawareness of information security.

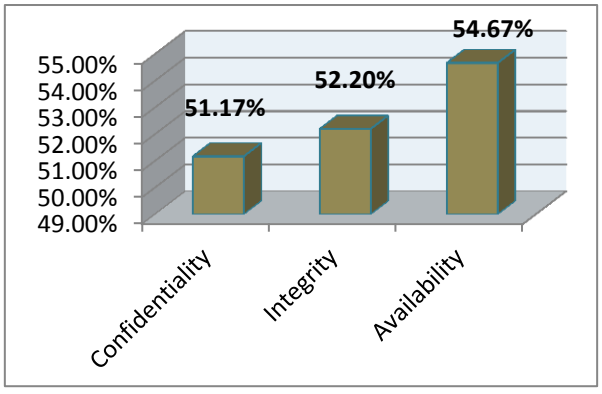
Fig 7: Bank(B) information security controls report



3) *High Level Risk Assessment of Bank (C)*

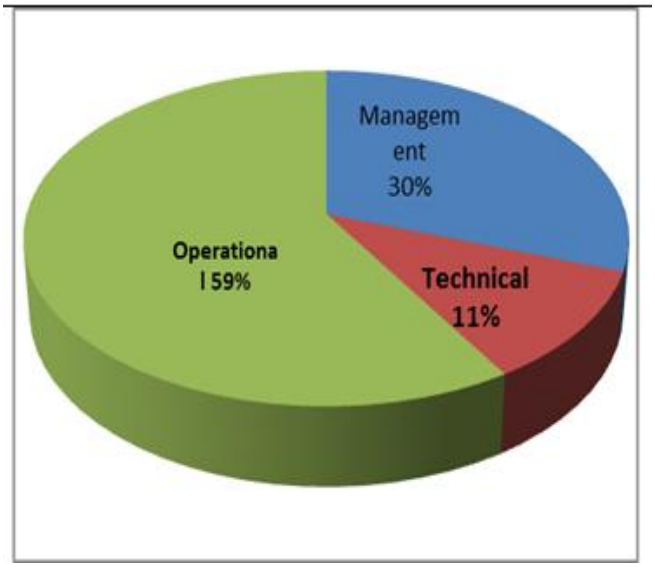
Bank (C) is the seventh largest bank of Pakistan with over 240 branches. The risk assessment done by COBRA™ is shown in Fig 8. The threats to availability of information were highest at 54.67% followed by integrity at 52.20% and confidentiality at 51.17%. These high threats showed that availability of critical information were to unauthorized and unwanted individuals or to the third-part. This could harm bank reputation and ultimately to its business. These high threats also showed vulnerability to correctness, completeness, and protection of information from intrusion.

Fig 8: Bank(C) information security risk report



The security controls bifurcation in bank (C) is shown in Fig 9. It may be seen that technical and management controls were lower being at 11% and 30%. Operational control maintained properly as its being at 59%. Risks associated to management controls like policy establishment and information security culture could be higher in this bank.

Fig 9: Bank(C) information security controls report

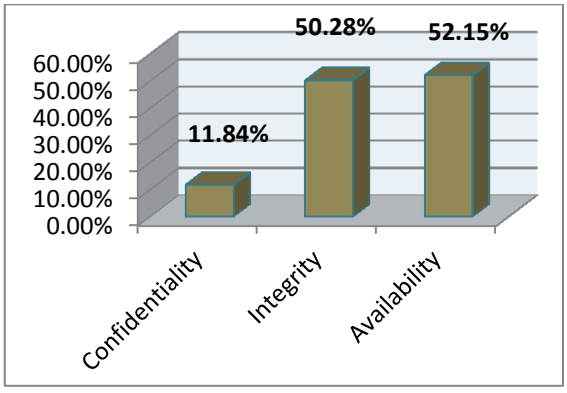


4) *High Level Risk Assessment of Bank (D)*

Bank (D) has the network of over 700 online branches in Pakistan.

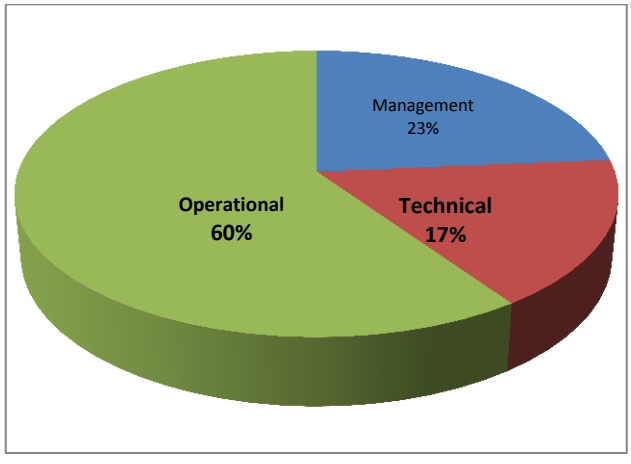
The risk assessment done by COBRA™ is shown in Fig 10. The threats to availability of information were highest at 52.15% followed by integrity at 50.28%. Threats to confidentiality of information were reported at 11.48%. The high threats of availability and integrity showed that critical information were accessible for unauthorized changes. On other hand information were not available at required time or were not managed properly.

Fig 10: Bank(D) information security risk report



The Fig 11 showed security controls percentage in bank (D). It may be seen that technical control is being at 17% than suggested 18%. Management control being at 23% is lower than suggested one. The operational control was being managed properly as it was at 69% against the suggested 31.76% by the software. The management control at 23% showed that a risk on policy establishment, weak information security culture, and poor management vision towards information security.

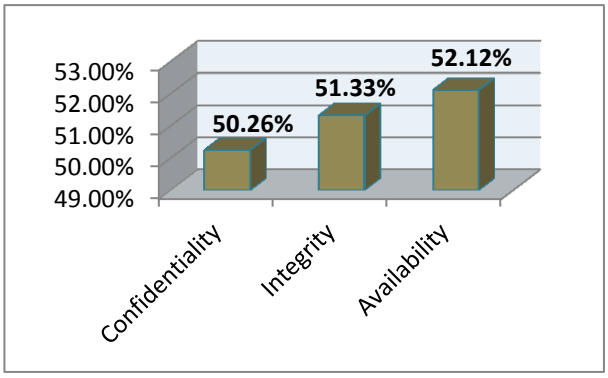
Fig 11: Bank(D) information security controls report



5) *High Level Risk Assessment of Bank (E)*

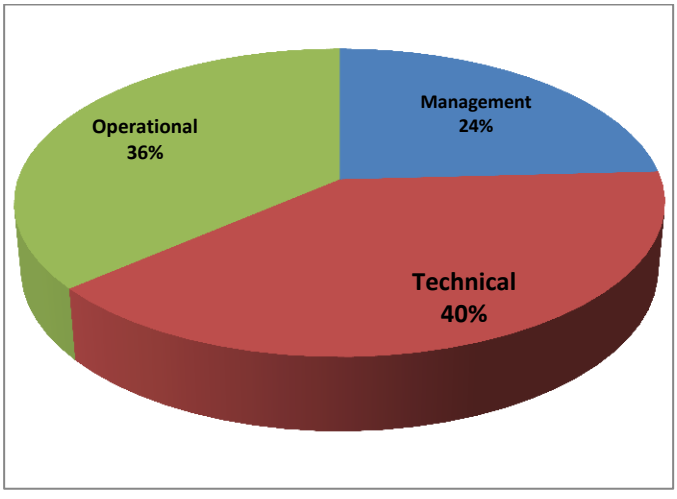
The Bank (E) provides microfinance services and act as a catalyst in stabilizing the country's newly formed microfinance sector. The risk assessment done by COBRA™ is shown in Fig 12. The threats to availability of information were highest at 52.12% followed by integrity at 51.33% and confidentiality at 50.26%. The high threats showed that critical information were accessible for unauthorized change. Availability of information was at required time or was not managed properly. Threat of unauthorized changes and completeness of information were also present.

Fig 12: Bank(E) information security risk report



The Fig 13 showed security controls percentage in bank (E). It may be seen that technical control was at 40% followed by operation control at 36% and management control at 24%. It is being observed from the Fig that management main focus is on technical control. Operational control is also maintained at 36% comparing with suggested 31.76% by software. The management control at 24% showed a risk on policy establishment, information security culture, and poor management vision towards information security

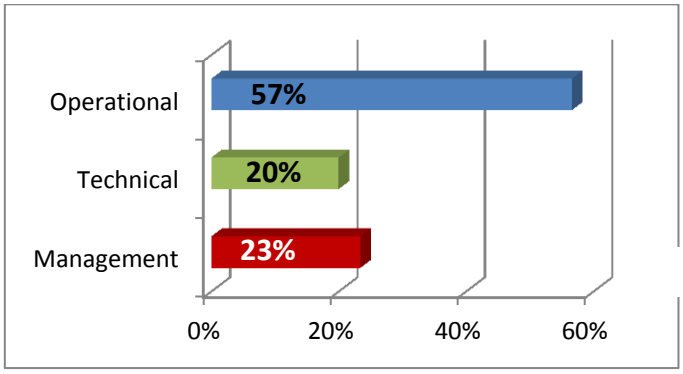
Fig 13: Bank(E) information security controls report



6) *Consolidated High Level Risk Assessment*

The security controls implemented in all banks is being evaluated in Fig 14. It is being seen that operational control in all banks was at 57% followed by management control at 23% and technical control at 20%.

Fig 14: Information security control in all banks



The COBRA™ suggested a percentage of 49.89% to management control, 31.76% to operational control, and 18.44% technical control for optimal information security. The comparative analysis of proposed and actual percentage of control showed that operational control in all banks was well maintained. The technical control was also maintained properly. The management control in all banks individually and in this consolidated report was at lowest percentage being at 23% shown that its not up-to the COBRA™ recommended mark that is almost 50%. Therefore the risk associated to management control must have to be high in all banks according to COBRA™ reports. Risk associated with the management control.

- Ineffective decision making
- Poor establishment of information security risk management policies/ procedures
- Unawareness of Information Security related risks
- Information secure culture
- Information Security not a part of overall business process.
- Fraudulent system usage
- Reputational damage
- Lack of business continuity planning
- Information security not a part of strategic planning

Consolidated Analysis of Management Contols

The second phase of survery is accomplished by developing sepecific questionnaire to evaluate the COBRA™ results. The questionnaire was divided into two sections. First section was about the management vision towards information security. Second section was about the information security awareness and information security culture in banks. A specific value was assigned to all the questions to have a quantitave analysis of banks management control.

To check the matuarity level of management control in these banks, overall score of the questinnaire was lied between -20 to 20. The maturity of management control was further classified into four levels shown in Fig 15.

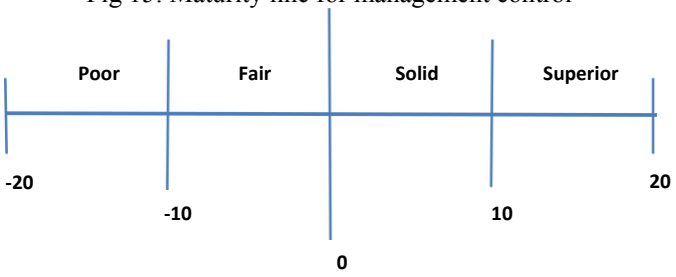
a. -20 to -10 = Poor

b. -10 to 0 = Fair

c. 0 to 10 = Solid

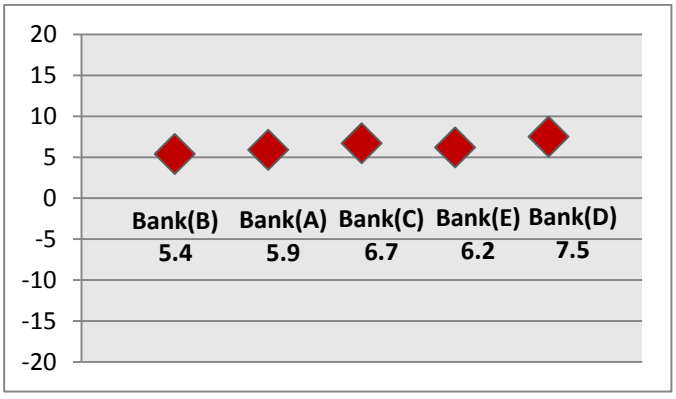
d. 10 to 20 = Superior

Fig 15: Maturity line for management control



Information security is more a management issue than a technical one. Effective management control is essential to establish information security culture. Establing secure information is a continus journey which can be achieved though action, policies, values, and positive management style. The questinnaires were filled by five personal of different management level of each bank. The result of all banks management control is shown in the Fig 16.

Fig 16: Maturity standing of all banks



The management controls in all banks is being at the solid level with a range from 5.4 to 7.5. The maximum management control is being in bank (D) at 7.5 and lowest in bank (B) at 5.4. The bank (A) is being at 5.9 followed by bank(C) and bank (E) at 6.7 and 6.2. Management control of all banks lied at the solid level. It was not in superior level in any of the banks.At solid level organization achieved the following management control:

- Information security policy is being rolled out
- Supporting standards and procedures are being developed
- Employee awareness has begun
- Confidentiality, Integrity, and availability of information is being considered
- Initial employee awareness process has begun
- Access to sensitive areas is generally restricted
- Employees are aware of fire safety procedures
- Contingency plans have been developed

- Management supports for information security has begun

VI. FINDINGS

COBRA™ gave a comprehensive information security risk analysis report but few short coming of this tool are:

- Risk raiting is not established properly in the COBRA™ e.g Fire cause more demage to information/ infrastructre etc than malfunctioning of any hardware. COBRA™ has given the same score to such cases.

- For risk assessment it is recommended to do the asset evalution of all the tangible and intangible assets of the organization. COBRA™ doest not evaluate individual asset value of organization in high level risk assessment. Due to this in case of any loss the accurate financial loss can not be predicted through this software.

- In risk assessment process, the range of accepted risks in COBRA™ is very low, from 0-19 score. any score above than 19 will be treated as a high expectancy of threat to organization. One drawback of this hard coded low risk acceptancy is that the risk level of all organization mostly falls in between 50% and above which in real scenario is exceptionally high risk for any organization. Secondly, perdiction for which organization like to use quantitaive tools than qualitative assessment tool is not obtainable.

- COBRA™ risk assessment reports coveres lot of information security risk area and also inform the requirement of security improvement at the exact areas, but do not inform the exact measures to mitigate them.

- Awareness of information security requirement to all employees of organization is essence of information security management system. Since as per NIST [9], employees are the biggest threat to organization information than any other attack. COBRA™ ignores that High Level Risk Assessment domain.

- Re-assessment of COBRA™ results by conducting second survey showed substantial differences, for instance, in COBRA™ reports the level of the management control implementation in all banks was between 13% to 30% whereas in re-assessment survey this range was between 50% to 75%.

- Besides all these drawbacks COBRA™ still facilitate the management in identification of information security risks.

VII. FUTURE WORK

Information security risk management framework which would cover the information security governance and show the results related to the information security controls so that organizations can focus and improve the deficient area regarding information security management.

VIII. REFERENCES

1) ISO/ IEC FDIS 17799 "Information Technology- security techniques- Code of practice for information security management" 2005.
2) ISO/ IEC FDIS 27001:2005(E) "Information Technology- Security techniques- Information Security Management Systems- Requirements", pp 1 – 9
3) Thomas Nowey and Hannes Federath "Collection of Quantitative Data on Security Incidents" 0-7695-2775-2/ 2007 IEEE.
4) Daniel Port, Rick Kazman, Ann Takenaka, Department of Information Technology Management, University of Hawaii "Strategic Planning for Information Security and Assurance" 978-0-7695-3126-7/ 2008 IEEE.
5) Fong-Hao Liu "Constructing Enterprise Information Network Security Risk Management Mechanism By Using Ontology" 0-7695-2847-3/ 2007 IEEE.
6) Ching- Jiang Chen and Ming-Hwa Li "SecConfig: A Pre-Active Information Security Protection Technique" 978-0-7695-3322-3/ 2008 IEEE
7) Heinz Lothar Grob, Gereon Strauch and Christian Buddendick" Applications for IT-Risk Management –Requirements and Practical Evaluation" DOI 0-7695-3102-4/ 2008 IEEE.
8) Wade H. Baker and Linda Wallace "Is Information Security Under Control? Investigating Quality in Information Security Management" 1540- 7993/ 2007 IEEE.
9) ENISA (European Network and Information Security Agency), "Risk Management: Implementation principles and Inventories for Risk Management/Risk Assessment method and tools", Available online: http://www.enisa.europa.eu, 2006 pp 39 - 51.
10) Julie J.C.H Ryan and Danel J. Ryan "Performance Metrics for Information Security Risk Management" 1540-7993/ 2008 IEEE.
11) Xiao Long, Qi Yong and Li Qianmu "Information Security Risk Assessment Based On Analytic Hierarchy Process and Fuzzy Comprehensive" 978-0-7695-3402-2/ 2008 IEEE.
12) Papadaki, K., Polemi, D., "Towards a systematic approach for improving information security risk management methods", in Proc. 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communication (PIMRC), 2007.

13) Thomas Finne, Abo Akademi University, Institute for Advance Management System Research(IAMSR) —A DSS for Information Security Analysis: Computer Support in a Company's Risk Management"0-7803-3280-6/ 1996 IEEE.
14) Symantec, —IT Risk Management Report 2: Myths and Realities", Available online at http://eval.symantec.com, 2008.
15) Brown, J S & Duguid, P, —Knowledge and organization: A social-practice perspective", Organization Science, 12, 2: 198-213, 2001.
16) Desouza, K.C.,Awazu, Y., Baloh, P. —Managing Knowledge in Global Software Development Efforts:
Issues and Practices", IEEE Software 23(5), 30–37, 2006.
17) Ekstedt M., et al., —Consistent Enterprise Software System Architecture for the CIO – A utility-Cost Approach", Proceedings of the 37th annual Hawaii International Conference on System Sciences (HICSS), 2004.
18) Johansson E., et al., —Assessment of EIS - An ATD Definition", in the Proceedings of the 3rd Annual Conference on Systems Engineering Research (CSER), March 23-25, 2005.
19) Johansson E., et al., —Assessment of Enterprise Information Security – The Importance of Prioritization", In the Proceedings of the 9th IEEE International Annual Enterprise Distributed Object Computing Conference (EDOC), Enschede, The Netherlands, September 19-23, 2005.
20) Edvardsson B., —The Need for Critical Thinking in Evaluation of Information", Proceedings of the 18th International Conference on Critical Thinking, Rohnert Park, USA, 1998.

Fuzzy Rule-based Framework for Effective Control of Profitability in a Paper Recycling Plant

Uduak A. Umoh[1], Enoch O. Nwachukwu[2], Obot E. Okure[1]

GJCST Classification
I.2.3 I.5.1

***Abstract*-The rapid and constant growth of urban population has led to a dramatic increase in urban solid waste production, with a crucial socio-economic and environmental impact. As the demand for materials continues to grow and the supply of natural resources continues to dwindle, recycling of materials has become more important in order to ensure sustainability. Recycling is one of the best ways for citizens to make a direct impact on the environment. Recycling reduces greenhouse gas emissions that may lead to global warming. Recycling also conserves the natural resources on Earth like plants, animals, minerals, fresh air and fresh water. Recycling saves space in the landfills for future generations of people. A sustainable future requires a high degree of recycling. Recycling industries face serious economic problems that increase the cost of recycling. This highlights the need of applying fuzzy logic models as one of the best techniques for effective control of profitability in paper recycling production to ensure profit maximization despite varying cost of production upon which ultimately profit, in an industry depend. Fuzzy logic has emerged as a tool to deal with uncertain, imprecise, partial truth or qualitative decision-making problems to achieve robustness, tractability, and low cost. In order to achieve our objective, a study of a knowledge based system for effective control of profitability in paper recycling is carried out. The root sum square of drawing inference is found to be the most suitable technique to infer data from the rules developed. This resulted in the establishment of some degrees of influence on the output. To reinforce the proposed approach, we apply it to a case study performed on Paper recycling industry in Nigeria. A computer simulation using the Matlab/Simulink and its Fuzzy Logic Tool Box is designed to assist the experimental decision for the best control action. The obtained simulation and implementation results are investigated and discussed.**

I. INTRODUCTION

Computational intelligence, the technical umbrella of Fuzzy Logic and Artificial Neural Networks (ANNs) have been recognized as powerful tools which is tolerant of imprecision and uncertainty and can facilitate the effective development of models by integrating several processing models. They explore alternative representation schemes, using, for instance, natural language, rules, semantic networks, or qualitative models. Fuzzy logic controllers have their origin with the E. H. Mamdani (Mamdani, 1977) researches, based on theories proposed by L. Zadeh (Zadeh, 1965). These controllers have founded space in many learning, research and development institutions around the world, being today an important application of fuzzy set theory.A great appeal of fuzzy technology in control is the possibility to operate with uncertainties and imprecision. It can consider an uncertainty on definition of input and output variables (Lee, 1995).The fuzzy controllers are robust and highly adaptable, incorporating knowledge that sometimes are not achieved by other systems. They are also versatile, mainly when the physical model is complex and has a hard mathematical representation. In fact, fuzzy controllers are especially useful in non-linear systems and plants with a high level noise. The conventional controllers deal with non-linearities of physical systems by approaching, considering the systems simply linearly, linear in parts, or describing them by extensive lookup tables that try to map the process inputs and outputs (Chin-Fan-Lin, 1994), (Cirstea, 2002). Fuzzy logic is a powerful technique for solving a wide range of industrial control and information processing applications (Akinyokun, 2002). Urbanization is one of the most evident global changes worldwide. The rapid and constant growth of urban population has led to a dramatic increase in urban solid waste production, with a crucial socio-economic and environmental impact. However, the growing concern for environmental issues and the need for sustainable development have moved the management of solid waste to the forefront of the public agenda.Recycling technology has evolved as one of the most useful facilities that help us to maintain an environmentally friendly society. All over the world, the need of recycling is heightened by the increasing awareness of product consumer, on the need to maximize the bundle of benefits from the products brought by them.
Consequently, the government is encouraging recycling and recycling practices (Goulias, 2001). For example, a strong indication of the —end-of-life directive" endorsed by the European Union states that if resources are not recycled, a period may come that very limited resources would be available for mankind (Treloar et al. 2003). If urgent steps are not taken, we may need to suffer for the shortages of these important resources. Waste paper is an example of a valuable material that can be recycled. Paper recycling has been around as long as paper itself. Paper companies have always recognized the environmental and economic benefits of recycling. Recycled paper reduces water pollution by 35%, reduces air pollution by 74%, and eliminates many toxic pollutants (TAPPI, 2001). In recent years, paper recycling has become popular with everyone as a way to help protect our environment by reusing our resources and conserving landfill space (TAPPI, 2001).Recycling is a series of activities, including collection, separation, and processing, by which products or other materials are recovered from or otherwise diverted from the solid waste stream for use in the form of raw materials in the manufacture of new products. Recycling is one of the best

About[1]- Department of Mathematics, Statistics, and Computer Science, University of Uyo, Nigeria
About[2]-Department of Computer Science, University of Port Harcourt, Nigeria.

ways for citizens to make a direct impact on the environment. Recycling reduces greenhouse gas emissions that may lead to global warming. Recycling also conserves the natural resources on Earth like plants, animals, minerals, fresh air and fresh water. Recycling saves space in the landfills for future generations of people. A sustainable future requires a high degree of recycling (Misman et al. 2008). Recycling industries face serious economic problems that increase the cost of recycling (Kumaran, 2001). By way of proffering solution, a number of management strategies are adopted over time. Such strategies may include business process re-engineering, downsizing, system restructuring, lean manufacturing, etc. All these strategies are aimed towards optimizing the system variables through minimization of costs and maximization of profits.The profitability of recycling industries has been known to be highly dependent on the effective management of resources and management practices, (Craighill and Powell, 1996) (Cunningham, 1969). Taking the paper recycling industry as a case study, the recycling of paper has contributed in no small measure to the conservation of material and consequent low cost of production.The ultimate resultant effect of the recycling process is high profit generated due to the low cost of production, hence, the link between profitability and recycling industry. Since the ultimate aim of any capitalist industry is to make profit, the concept of profitability is of great significance in science and engineering (http://www.recyclingtoday.com, http://www.recyclinginternational.com,).The objective of this research is to develop a fuzzy logic model using fuzzy logic technology and apply the model for effective control of profitability in paper recycling to determine the degree of influence of various processing components on the profit generated by an industry. We derive the cost of production from the following costs; (i) cost incurred to recycle the desired quantity of waste paper materials (secondary fiber), (ii) salaries, (iii) wages, (iv) rent, (v) depreciation on machines and equipment, (vi) sundry expenses, etc. We determine the selling price based on the production cost and then generate profit based on the relationship between the above mentioned functions. A relationship is developed between thestated input(s) (cost of production and selling price) and output(s) (profits) with the help of the fuzzy logic model. This will facilitate optimization of paper recycling production.The proposed model will help to arrive at specified desired output in the optimization approach for solid waste paper recycling. To reinforce the proposed approach, we have applied it to a case study performed on Paper recycling industry in Nigeria. A computer simulation using the Matlab®/Simulink and its Fuzzy Logic Tool Box is designed to assist the experimental decision for the best control action. The obtained simulation and implementation results are investigated and discussed.Matlab is an integrated technical computing environment that combines numerical computation, advanced graphics visualization and a high-level programming language, Simulink is built on top of Matlab, and is an interactive environment for modeling, analyzing and simulating a wide variety of dynamic systems. The Fuzzy Logic Toolbox provides tools for user to create and edit fuzzy inference systems, or integrate the fuzzy systems into simulations with Simulink. (Wongthatsanekorn, 2009) developed a goal programming model for plastic recycling system in Thailand. (Kufman, 2004) carried out the analysis of technology and infrastructure of paper recycling. (Kumar et al, 2008) designed a goal programming model for paper recycling to assist proper management of the paper recycling logistic system, while (Udoakpan, 2002) carried out the financial implication of establishing a paper recycling plant in Nigeria. (Oke et al, 2006) designed a fuzzy logic model to handle the profitability concept in a plastic industry.In section 2 of the paper the mathematical model of the system is presented while in Section 3 the research methodology is presented. Section 4 presents the model experiment while in Section 5 results of findings are discussed. Finally in Section 6, some recommendations are made and conclusion is drawn.

II. MATHEMATICAL MODEL

The mathematical model of the proposed work is based on the major components in the concept of profitability using a case in the paper recycling industry. These components include; Selling Price (SP), Cost Price (CP), Quantity Recycled (QR), and Profit (Y) The relationship among these components as used in the concept of profitability is illustrated in figure 1.The relationships among components of profitability in figure 1 show Selling Price (SP) which is the selling price per item of the quantity of recycled product. Cost Price (CP) which is the cost price per item of the quantity of recycled product. The cost price is made up of all expenses incurred directly during the recycling production processes

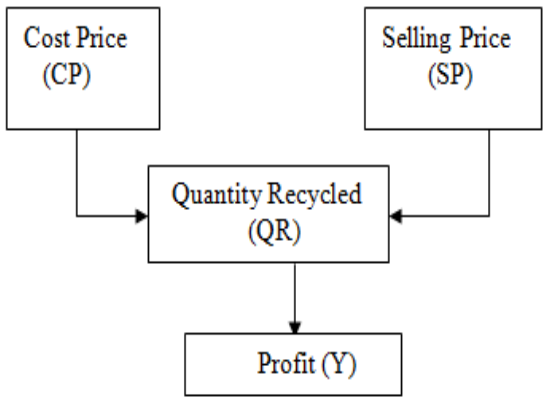


Figure 1: Relationships among Components of Profitability (adopted from Oke et al, 2006)

While Quantity Recycled (QR) is the quantity of recycled product which is the output from wastepaper input and it is determined by the conventional control model, based on the following parameters of the paper recycling process:

PM4 = Machine Speed (m/m)

T = Period between the beginning of recycling and the expected completion time.

SUBW = Width of the substance (m)

SUBP = Paper substance or grammage (g/m^2)
STRNG = Strange (%) which is the difference between the machine speed and the reelers speed
Thus,
QR = PM4 (m/m) x T (m) x SUBW (m) x SUBP (g/m^2) x STRNG (%) (1)
Profit (Y) is the profit made and is the difference in Selling Price (SP) and Cost Price (CP) multiplied by the quantity of recycled product (QR). This is given as:
Y = SP (QR) – CP (QR) (2)
Thus,
Y = [(SP – CP) (PM4 (m/m) x T (m) x SUBW (m) x SUBP (g/m^2) x STRNG (%))] (3)

III. Research Methodology

The fuzzy inference system for effective control of profitability in paper recycling is shown in Figure 2. This system involves three main processes; fuzzification, inference and defuzzification. The knowledge base contains the following:
(i) rule-base - that contains knowledge used to characterize Fuzzy Control Rules and Fuzzy Data Manipulation in an FLC, which are defined based on experience and engineering judgment of an expert. In this case, an appropriate choice of the membership functions of a fuzzy set plays a crucial role in the success of an application. The rules are in the form of IF – THEN (production rules).
(ii) data-base: Fuzzy variables are defined by fuzzy sets, which in turn are defined by membership functions. The knowledge base design of profitability control in paper recycling production is made up of both static and dynamic information about the decision variables and about the different factors that influence recycling decision for controlling paper recycling production for profit optimization. There are qualitative and quantitative variables which must be fuzzified , inferred and defuzzified.
Fuzzification of data is carried out on the transformed data by selecting input parameters into the horizontal axis and projecting vertically to the upper boundary of membership function to determine the degree of membership. This is then used to map the output value specified in the individual rules to an intermediate output measuring fuzzy sets. Parameters used in fuzzy logic model are, Cost Price (CP), Selling Price (SP), and Quantity of Recycled product (QR). These parameters constitute the fuzzy logic input variables used to generate the fuzzy logic model, while the linguistic variable for the model is (SP)(QR) – (CP)(QR) which is the difference between the selling price and cost price of the quantity recycled.

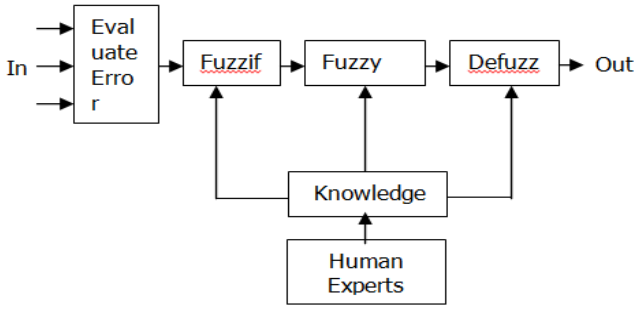


Figure 2: Empirical Fuzzy Logic Model for Control of Profitability in Paper Recycling Production

The following error and change in error terms of the model as defined by (Oke et al, 2006)and are modified and evaluated thus.
Error (E) = (SP – CP)(QR)
Change inn error (CE) = (ΔSP - ΔCP)(QR) = d/dt Error
Error equals selling price of the quantity recycled minus cost price of the quantity recycled. Change in error equals differentiating selling price of quantity recycled minus cost price of quantity recycled over time.

1) *Error terms*

(SP)(QR) – (CP)(QR) = ZE, "Zero-error" term(ZE) (No profit no loss)
$(S_P)(Q_R)$ – $(C_P)(Q_R)$ = NE, "Negative-error" term (NE) (Loss)
$(S_P)(Q_R)$ – $(C_P)(Q_R)$ = PO, "Positive-error" term (PO) (Profit)
If we consider the model over a period of time, we have:

2) *Change in Error Terms*

d{(SP)(QR)-(CP)(QR)}/dt = Z, "Zero error-change" (ZE) (No profit no loss over time)
d{(SP)(QR)-(CP)(QR)}/dt = N, "Negative error-change" (NE) (Loss over time)
d{(SP)(QR)-(CP)(QR)}/dt = P, "Positive error-change" (PO) (Profit over time)
We employ the characteristic fuzziness of the model by generating more "error" and "change in error terms by considering the model as changing or varying to large degree over time in order to achieve more effective control of the fuzzy logic model.

3) *More error terms*

(SP)(QR) - (CP)(QR) = << N, "Negative Small error" (NS) (Low Loss)
(SP)(QR) - (CP)(QR) = >>P, "Positive Small error" (PS) (Low Profit)
(SP)(QR) - (CP)(QR) = <<<< N, "Negative Big error" (NB) (High Loss)
(SP)(QR) - (CP)(QR) = >>>> P, "Positive Big error" (PB) (High Profit)

4) *More change in error terms*

d{(SP)(QR)-(CP)(QP)}/dt = << N, —Negative Small error-change" (NS) (Low loss over time)
d{(SP)(QR)-(CP)(QP)}/dt = >> P, —Positive Small error-change" (PS) (Low profit over time)
d{(SP)(QR)-(CP)(QR)}/dt = <<<< N, —Negative Big error-change" (NB) (High Loss over time)
d{(SP)(QR)-(CP)(QR)}/dt = >>>> P, —Positive Big error-change" (PB) (High Profit over time)

In this paper, the universes of discourse for error (E),change in error (CE) and Output are chosen to be [-100, 100], [-100, 100], and [-100, 100] respectively. Both sets of the linguistic values for the linguistic variables E and CE are {NB, NS, N, Z, P, PS, PB}, and the set of linguistic values for Output is {HL, LL, L, NPNL, P, LP, HP}, where NB, NS, N, Z, P, PS, and PB represent negative_big, negative_small, negative, zero, positive,

$$\mu(x) = \begin{cases} 0 & \text{if } x < a_1 \\ x - a_1/a_2 - a_1 & \text{if } a_1 <= x < a_2 \quad (4) \\ a_3 - x/\ a_3 - a_2 & \text{if } a_2 <= x < a_3 \\ 0 & \text{if } x > a_3 \end{cases}$$

$$\text{Error}(x) = \begin{cases} 0 & \text{”NB”} \\ (x+63)/34 & \text{if } -67 <= x < -33 \quad \text{“NS”} \\ - x/33 & \text{if } -33 <= x < 0 \quad \text{“NE”} \\ (33-x)/66 & \text{if } 0 <= x < 33 \quad \text{“ZE”} \\ (67-x)/34 & \text{If } 33 <= x < 67 \quad \text{“PO ”} \\ (100-x)/33 & \text{If } 67 <= x < 100 \quad \text{“PS”} \\ 0 & \text{if } x >= 100 \quad \text{“PB”} \end{cases} \quad (5)$$

$$\mu_{NS}(x) = \begin{cases} 0 & \text{if } x <= -100 \\ (x + 100)/33 & \text{if } -100 <= x < -67 \\ (-33+x)/34 & \text{if } -67 <= x < -33 \\ 0 & \text{if } x >= -33 \end{cases} \quad (6)$$

$$\mu_{NE}(x) = \begin{cases} 0 & \text{if } x < -67 \\ (x+67)/34 & \text{if } -67 <= x < -33 \\ -x/33 & \text{if } -33 <= x < 0 \\ 0 & \text{if } x >= 0 \end{cases} \quad (7)$$

$$\mu_{ZE}(x) = \begin{cases} 0 & \text{if } x < -33 \\ (x + 33)/33 & \text{if } -33 <= x < 0 \quad (8) \\ (33 - x)/33 & \text{if } 0 <= x < 33 \\ 0 & \text{if } x > 33 \end{cases}$$

$$\mu_{PS}(x) = \begin{cases} 0 & x <= 33 \\ (x-33)/34 & \text{if } 33 <= x < 67 \\ (100-x)/33 & \text{if } 67 <= x < 100 \quad (10) \\ 0 & \text{if } x >= 100 \end{cases}$$

$$\mu_{PB}(x) = \begin{cases} 0 & \text{if } x < 67 \\ (x-67)/33 & \text{if } 67 <= x < 100 \quad (11) \\ 1 & \text{if } x >= 100 \end{cases}$$

$$\mu_{HLoss}(x) = \begin{cases} 0 & \text{if } x <= -100 \\ (x + 100)/25 & \text{if } -100 <= x < -75 \\ (-50-x)/25 & \text{if } -75 <= x < -50 \quad (12) \\ 0 & \text{if } x >= -50 \end{cases}$$

$$\mu_{LLoss}(x) = \begin{cases} 0 & \text{if } x < -75 \\ (x+75)/25 & \text{if } -75 <= x < -50 \\ (-25-x)/25 & \text{if } -50 <= x < 25 \quad (13) \\ 0 & \text{if } x >= 25 \end{cases}$$

$$\mu_{Loss}(x) = \begin{cases} 0 & \text{if } x < -50 \\ (x + 50)/25 & \text{if } -50 <= x < -25 \\ - x/25 & \text{if } -25 <= x < 0 \quad (14) \\ 0 & \text{if } x > 0 \end{cases}$$

$$\mu_{NPNL}(x) = \begin{cases} 0 & \text{if } x <= -25 \\ (x+25)/25 & \text{if } -25 <= x < 0 \\ (25-x)/25 & \text{if } 0 <= x < 25 \quad (15) \\ 0 & \text{If } x > 25 \end{cases}$$

$$\mu_{Profit}(x) = \begin{cases} 0 & \text{if } x <= 0 \\ x/25 & \text{if } 0 <= x < 25 \\ (50-x)/25 & \text{if } 25 <= x < 50 \quad (16) \\ 0 & \text{if } x >= 50 \end{cases}$$

$$\mu_{LProfit}(x) = \begin{cases} 0 & \text{if } x < 25 \\ (x-25)/25 & \text{if } 25 <= x < 50 \\ (25-x)/25 & \text{if } 50 <= x < 75 \\ 0 & \text{if } x > 75 \end{cases} \quad (17)$$

$$\mu_{HProfit}(x) = \begin{cases} 0 & \text{if } x < 50 \\ (x-50)/25 & \text{if } 50 <= x < 75 \\ (100-x)/25 & \text{if } 75 <= x < 100 \\ 0 & \text{if } x > 100 \end{cases} \quad (18)$$

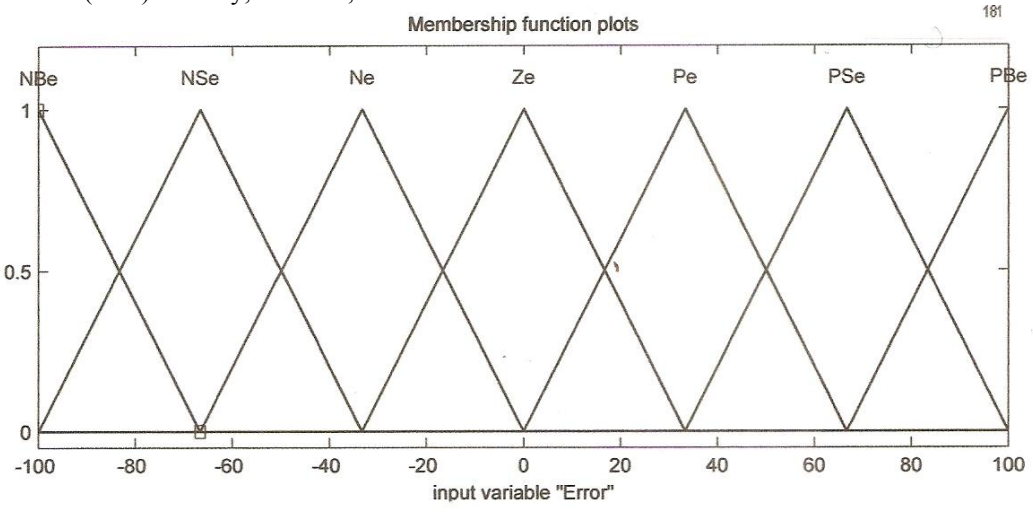
positive_small, and positive_big, respectively. The output is defined by fuzzy sets, high_loss, low_loss, loss, no_profit_no_loss, profit, low_profit and high_profit.The linguistic expression E and CE variables and their membership functions are evaluated using triangular membership function as presented in equations (5) to (11). Triangular curves depend on three parameters a_1, a_2, and $a_3$and are given by equation (4), a_2 defines the triangular peak location, while a_1 and a_3 define the triangular end points. During the process linguistic labels (values) are assigned to the error and change in error indicating the associated degree of influence of membership for each linguistic term that applies to that input variable.

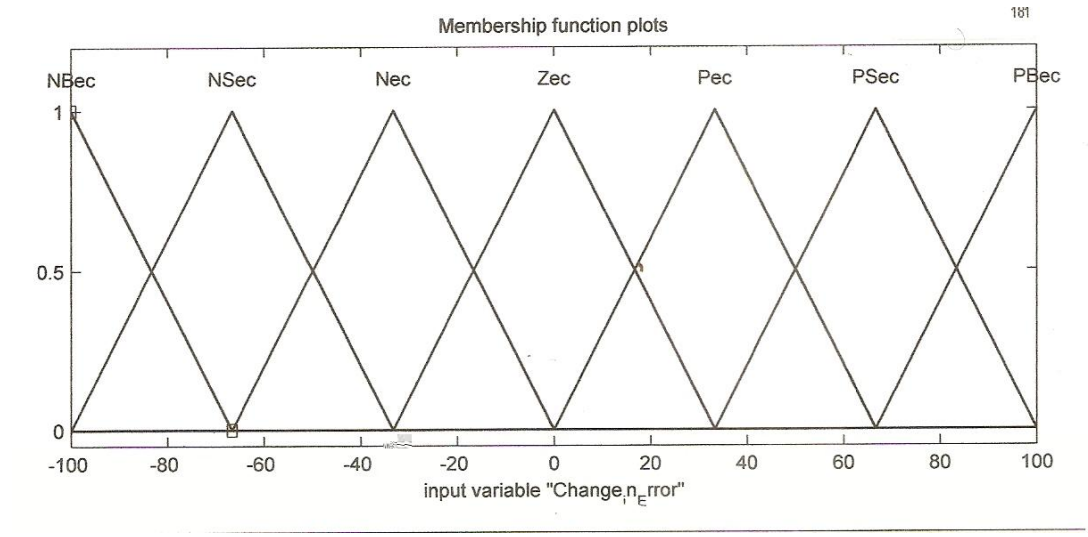
Degrees of membership (U_x) are assigned to each linguistic value as expressed in equations (5) to (11) negative small, negative, zero, positive, positive small and positive big.

Linguistic values are assigned to the linguistic variable, error ((SP)(QR) – (CP)(QR)) of profitability as shown in equation (5). In equations (6) to (11), each linguistic value is assigned a label emphasizing the degree of the value assigned in (1). For example, equation (6) evaluates the degree of positive small of the error and change in error, if the value of error is for instance, 55, the degree of influence will evaluate to 0.65 (65%) severity, whereas, 75 evaluates to 0.75 (75%). Fuzzy logic toolbox in Matlab 2007 is employed in this project to model the design. Graphical users interface (GUI) tools are provided by fuzzy logic box (Fuzzy Logic Toolbox Users‘ guide, 2007). The graphical formats which show the fuzzy membership curves for error, change in error and the output are depicted in figures 3, 4 and 5 respectively, where triangular membership functions are used to describe the variables.

linguistic value of fuzzy output membership function in figure 5 is assigned a label emphasizing the degree of the value assigned as in equations 12-.18

Using derivation based on expert experience and control engineering knowledge; the experience of an expert who has been working at the Star Paper Mill located in Aba, Nigeria for over 18 years was used to obtain the rule base. The expert also assisted in defining the fuzzy rules and the fuzzy set. There are 2 inputs in the knowledge base namely; error and change in error, with 7 fuzzy sets each as antecedent parameters and 7 fuzzy sets each as consequent parameters. From the expert knowledge, these are used to generate 49 rules for the rule base defined for the decision-making unit. Some of the rules are presented in Table 1. The Rule matrix for the fuzzy control rules is shown in Table 2





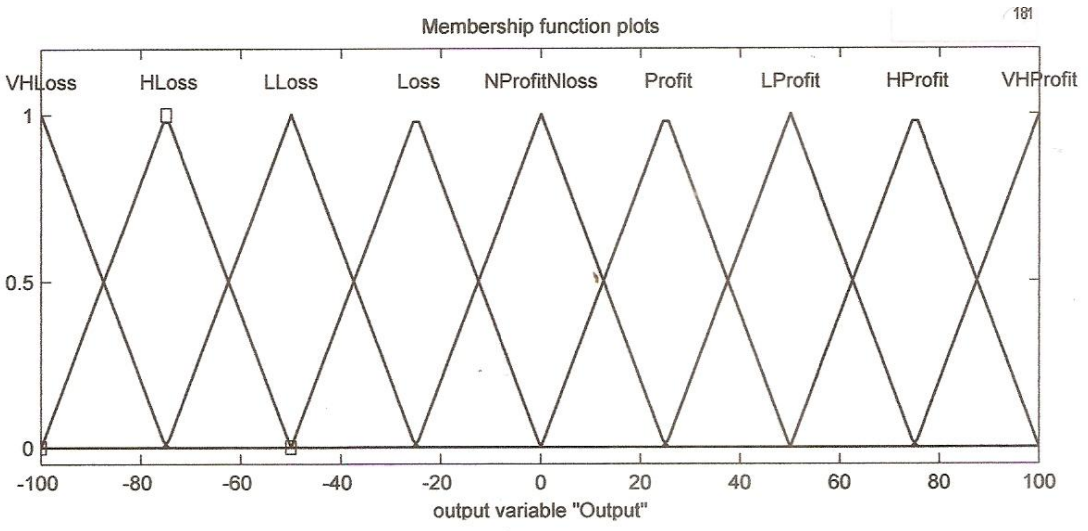


Table 1 Fuzzy rule base

Rule No.	Rules
1.	IF $(S_P)(Q_R)$-$(C_P)(Q_R)$ = NB AND d[$(S_P)(Q_R)$- $(C_P)(Q_R)$]/dt = NB THEN Output = HighLoss
4.	IF $(S_P)(Q_R)$-$(C_P)(Q_R)$ = NB AND d[$(S_P)(Q_R)$- $(C_P)(Q_R)$]/dt = ZETHEN Output = NProfitNLoss
8.	IF $(S_P)(Q_R)$-$(C_P)(Q_R)$ = NS AND d[$(S_P)(Q_R)$- $(C_P)(Q_R)$]/dt = NB THEN Output = HighLoss
14.	IF $(S_P)(Q_R)$-$(C_P)(Q_R)$ = NS AND d[$(S_P)(Q_R)$- $(C_P)(Q_R)$]/dt = PB THEN Output = Profit
18.	IF $(S_P)(Q_R)$-$(C_P)(Q_R)$ = NE AND d[$(S_P)(Q_R)$- $(C_P)(Q_R)$]/dt = ZETHEN Output = NProfitNLoss
19.	IF $(S_P)(Q_R)$-$(C_P)(Q_R)$ = NE AND d[$(S_P)(Q_R)$- $(C_P)(Q_R)$]/dt = PO THEN Output = LowProfit
23.	IF $(S_P)(Q_R)$-$(C_P)(Q_R)$ = ZE AND d[$(S_P)(Q_R)$- $(C_P)(Q_R)$]/dt = NS THEN Output = LowLoss
24.	IF $(S_P)(Q_R)$-$(C_P)(Q_R)$ = ZE AND d[$(S_P)(Q_R)$- $(C_P)(Q_R)$]/dt = NE THEN Output = Loss
28.	IF $(S_P)(Q_R)$-$(C_P)(Q_R)$ = ZE AND d[$(S_P)(Q_R)$- $(C_P)(Q_R)$]/dt = PB THEN Output = HighProfit

Table 2: Fuzzy control rules matrix

E	NB	NS	N	Z	P	PS	PB
NB	HL	HL	HL	NPNL	LP	LL	P
NS	HL	LL	L	NPNL	LL	NPNL	P
N	HL	LL	L	NPNL	LL	P	LP
Z	HL	LL	L	NPNL	LP	LP	HP
P	LL	L	NPNL	P	P	LP	HP
PS	L	LP	P	LP	LP	LP	HP
PB	NPNL	P	LP	LP	HP	P	HP

5) *Fuzzy Inference*

The process of drawing conclusions from existing data is called inference. For each rule, the inference mechanism looks up the membership values in the condition of the rule. Fuzzy input are taken to determine the degree to which they belong to each of the appropriate fuzzy sets via membership functions. The aggregationoperation is used to calculate the degree of fulfillment or firing strength, α_n of the condition of a rule n. A rule, say rule 1, will generate a fuzzy membership value μ_{E1} coming from the errorand a membership value μ_{EC1} coming from the change in error measurement. μ_{E1} and μ_{EC1} are combined by applying fuzzy logical AND to evaluate the composite firing strength of the rule. The rules use the input membership values as weighting factors to determine their influence on the fuzzy output sets of the final output conclusion. The degrees of truths (R) of the rules are determined for each rule by evaluating the nonzero minimum values using the AND operator. Only the rules that get strength higher than 0, would ―fire‖ the output. The Root Sum Square (RSS) inference engine is employed in this research which has the formula,

$$RSS = \sqrt{\sum_R}^2 = \sqrt{(R_1^2 + R_2^2 + R_3^2 + ,\ldots, + Rn^2)} \quad (19)$$

Where R_1, R_2, R_3…. R_n are strength values of different rules which share the same conclusion. RSS method combines the effects of all applicable rules, scales the functions at their respective magnitudes, and computes the "fuzzy" centroid of the composite area. This method is more complicated mathematically than other methods, but is selected for this work since it gives the best weighted influence to all firing rules (Saritas and Sert 2003). From table 1 for instance, the membership function strength values are evaluated as,

$$HighProfit = \sqrt{(R_{28}^2 + R_{42}^2 + R_{49}^2)} \quad (20)$$

$$LowProfit = \sqrt{(R_{19}^2 + R_{37}^2 + R_{41}^2 R_{45}^2)} \quad (21)$$

$$Loss = \sqrt{(R_{24}^2 + R_{36}^2)} \quad (22)$$

6) *Defuzzification*

Defuzzification of data into a crisp output is a process of selecting a representative element from the fuzzy output inferred from the fuzzy control algorithm. A fuzzy inference system maps an input vector to a crisp output value. In order to obtain a crisp output, we need a defuzzification process. The input to the defuzzification process is a fuzzy set (the aggregated output fuzzy set), and the output of the defuzzification process is a single number. Many defuzzification techniques are proposed and four common defuzzification methods are center-of-area (gravity), center-of-sums, max-criterion and mean of maxima. According to (Obot, 2008), max-criterion produces the point at which the possibility distribution of the action reaches a maximum value and it is the simplest to implement. The center of area (gravity)is the most widely used technique because, when it is used, the defuzzified values tend to move smoothly around the output fuzzy region, thus giving a more accurate representation of fuzzy set of any shape (Cochran and Chen, 2005). The technique is unique, however, and not easy to implement computationally. Center of gravity (CoG) often uses discretized variables so that CoG, y′ can be approximated to overcome its disadvantage as shown in equation (23) which uses weighted average of the centers of the fuzzy set instead of integration. This approach is adopted in this research because it is computationally simple and intuitively plausible.

$$y' = \frac{\sum \mu(x_i)\, x_i}{\sum \mu(x_i)} \quad (23)$$

Where x_i is a running point in a discrete universe, and $\mu(x_i)$ is its membership value in the membership function. The expression can be interpreted as the weighted average of the elements in the support set.

IV. Model Experiment

The study adopts Matlab/Simulink and its Fuzzy Logic tool box functions to develop a computer simulation showing the user interface and fuzzy inference to assist the experimental decision for the best control action. Results of evaluation of fuzzy rule base inference for two ranges of inputs, Error ((SP)(QR) – (CP)(QR)) and Change in error (d{(SP)(QR) – (CP)(QR)}/dt) are shown in Tables 3 and Table 4 respectively.

For example, rules 18, 19, 25, 26, 32 and 33 fire (generate non-zero output conclusions) from the rule base in figure 1 when error and change in error are selected at -18 and +18, their corresponding —Z membership = 0.5 and —N" membership = 0.5. A —Z" and "P" membership degree of 0.5 is indicated for change in error. An

Table 3 Rule base evaluation for error and change in error at -18 and +18

error of -18 and change in error of +18

selects regions of the "no profit no loss", —profit" and "low profit" output membership functions. The respective output membership function strengths (range: 0-1) from the possible rules (R1-R49) are computed using RSS inference technique as follows

$$Z = \sqrt{(R_{18}^2 + R_{25}^2)} = \sqrt{((0.5)^2 + 0.5)^2)} \quad (24) \quad =$$

Rule No.	Premise Variables		Conclusion Part of rule	Minimum value (non zero)
	Error	Change in error		
18	0.5	0.5	NoProfitNoLoss	0.5
19	0.5	0.5	LowProfit	0.5
25	0.5	0.5	NoProfitNoLoss	0.5
26	0.5	0.5	LowProfit	0.5
32	0.5	0.5	Profit	0.5
33	0.5	0.5	Profit	0.5

0.707 (NoProfitNoLoss)

$$P = \sqrt{(R_{32}^2 + R_{33}^2)} = \sqrt{((0.5)^2 + (0.5)^2)} = 0.707$$ (Profit)

$$PS = \sqrt{(R_{32}^2 + R_{33}^2)} = \sqrt{((0.5)^2 + 0.5)^2)} = 0.707$$ (Low Profit)

This is then defuzzified to obtain the crisp output for the above range.

$$\frac{((-18 \times 0.00) + (0 \times 0.707) + (25 \times 0.707) + (50 \times 0.707))}{+ 0.707)} = 25.7\% \text{ Profit}$$ (25)

These particular input conditions indicate positive value of 25.7% (25.7% Profit) therefore profit is expected with 25.7% possibility and required system response.

Table 4 Rule base evaluation for error and change in error at +95 and +95

Table 4 shows that, if rules 41, 42, 48 and 49 fire from the rule base in figure 1, when error and change in error are selected at +95 and +95 and their corresponding degrees of membership are "PS" = 0.2, "PB" = 0.8 and 0.0 in other fuzzy sets for error and "PS" = 0.2, "PB" = 0.8 and 0.0 in other fuzzy sets for change in error, the Root Sum Square inference for Profit (P), Low Profit (LP) and High profit

Rule No.	Premise Variables		Conclusion Part of rule	Minimum value (non zero)
	Error	Change in error		
41	0.2	0.2	LowProfit	0.2
42	0.2	0.8	HighProfit	0.2
48	0.8	0.2	Profit	0.2
49	0.8	0.8	HighProfit	0.8

(HP) membership functions is calculated as follows

$P = \sqrt{R_{48}^2} = \sqrt{(0.2)^2} = 0.2$ (Profit)

$PS = \sqrt{R412} = \sqrt{(0.2)2} = 0.2$ (Low Profit)

$PB = \sqrt{(R_{42}^2 + R_{49}^2)} = \sqrt{((0.2)^2 + (0.2)^2)} = \sqrt{(0.4 + 0.64)} = 1.0198$ (High Profit) (26)

$$\text{Crisp output} = \frac{(25 \times 0.2) + (50 \times 0.2) + (75 \times 1.0198)}{0.2 + 0.2 + 1.0198}$$

$$= 64.4\% \text{ (LProfit)} \quad (27)$$

The centers of the triangles representing the NB, NS, N, P, PS and PB membership functions for the two inputs are manipulated so as to achieve the desired result

The values of the errors and change in errors indicated as in Tables 3 and 4 are inserted into the rule base under the view rule editor and the outputs computed for all the cases are recorded. The inference mechanism of fuzzy sets in our 2 examples is shown in figures 6 and 7 (generated in the Matlab Fuzzy Logic Toolbox)

V. Result and Discussion

In fuzzy logic implementation, the selection of membership functions and rule base determine the output. Hence, by selecting a triangular membership function, the variables in the system are manipulated and represented judiciously. Also, the rule base is selected from the experience of system expert. Fuzzy logic represents partial —truth" or partial "false" in its modeling. From the study, apart from assigning linguistic variables such as low-loss, no-profit-no-loss, low-profit,to the profitability, the degree of influence or severity of each linguistic variable is evaluated. Tables 3 and 4 show fuzzy logic model of the variables, error (E) and change in error (CE) in order to remove uncertainty, ambiguity and

vagueness. The crisp outputs in the two examples cited in this work show the linguistic label and degree of influence on profitability. From the crisp outputs obtained from the graph of fuzzy logic model in figures 6 and 7, it is observed that these particular input conditions indicate positive values of 26.3% (26.3% Profit) and 63.7% (63.7% Low-profit). This implies that the selling price of the recycled quantity is more than the cost price of the recycled quantity; therefore profit is expected with 25.7% possibility and a low-profit with 63.7% possibility. Considering the degree of relationship between linguistic label and value of fuzzy output membership function, say "Loss", when its value equals 1.0, it indicates that the cost price of the recycled quantity is more than the selling price of the recycled quantity and that industry will run at a loss with 100% possibility. When the fuzzy output value is 0.6, it indicates 60% possibility of loss.Considering the relationships strength among fuzzy outputs in figure 5, it indicates that only when "no-profit-no-loss" output value equals 1.0, (100%) that we can conclude that the selling price of the quantity recycled is indeed equal to the cost price and the industry is likely to run at no profit no loss. Relating "no-profit-no-loss' with "profit" for instance, when the value of "noprofit-loss" output is 0.4 showing possibility 40%, its indicates that there is 0.6 (60%) possibility of profit. This implies that it is not likely that the industry will run at no profit no loss altogether when the selling price is only less than or equal to the cost price by 40%. Relating "no-profit-no-loss" with "loss" with the relationship strength of 0.5 (50%), it shows no profit no loss with 0.5 (50%) possibility and 0.5 (50%) of loss in this case.

Several responses can be observed during the simulation of the system. The system is tuned by modifying the rules and membership functions until the desired system response (output) is achieved. The system can be interfaced to the real world via Java programming language.

Figure 6: Graphical construction of the inference mechanism of fuzzy sets in example 1

7: Graphical construction of the inference mechanism of fuzzy sets in example 2

VI. Conclusion

It is important to make evident the great potential that fuzzy logic has to offer, such as the need for the mathematical model. Fuzzy Logic Controllers can provide more effective control of non-linear systems than linear controllers, as there is more flexibility in designing the mapping from the input to the output space. Fuzzy logic is capable of resolving conflicts by collaboration, propagation and aggregation and can mimic humanlike reasoning. Another advantage that the fuzzy logic offers is that an autotuning algorithm can be applied to the system, by the means of this reasoning. In this way, the system can learn the control parameters to take. In our study, we represent the mathematical expression profitability components using linguistic variables. We consider —error" and —change in error" approaches in a vague, ambiguous and uncertain situation. It is shown that fuzzy logic is able to represent common sense knowledge and address the issue of vagueness, ambiguity and uncertainty (Obot, 2008) as it is used to find the exact degree of profit, loss, low-profit, etc in the profitability of an industry. To this end, fuzzy logic can be used to control and ensure the desired output in a model since it can tolerate wide variation in input variables. Fuzzy logic control model shows that profit can be achieved at various levels, but maximum profit is achieved when the selling price is more than the cost price by 100% (1.0). Also loss can be incurred at different levels when the production cost is more than the selling price. The exact level and exact loss or profit has been clearly defined by fuzzy logic control system thereby resolving the conflict of uncertainty and vagueness. Our case study reveals that production cost depends on waste paper cost, the parameters of paper recycling process and other costs associated with paper recycling. Therefore it is recommended that the parameters affecting production cost, and determine the output of paper recycling be modified so as to reduce production cost and achieve maximum profit. Since the ultimate aim of any capitalist industry is to make profit, the concept of profitability is of great significance and it is evident that the fuzzy logic model developed, if implemented is an effective tool to effectively control profitability in paper recycling to achieve maximum profit. For further optimization of results of our work, hybridization of fuzzy logic and neural network or fuzzy logic and genetic algorithm is recommended for future research.

VII. REFERENCES

1) Akinyokun, o. c. (2002). Neuro- Fuzzy Expert System for Evaluation of Human Resources Performance. First Bank of Nigeria PLC Endowment Fund Lecture Series 1, Delivered at the Federal University of Technology, Akure, December, 10, 2002.
2) Cochran, J. K. and Chen, H. (2005). Fuzzy Multi-criteria Selection of Object-Oriented Simulation Software for production System Analysis. Computer and operation Research, Vol. 32, pp153-168.
3) Chin-Fan-Lin. (1994), Fuzzy Logic Controller Design". Advanced Control Systems Design.Prentice Hall: Englewood Cliffs, NJ. 431-460.
4) Cirstea, M. N., Dinu, A. Khor, J. G and McCormick, M. (2002), Neural and Fuzzy Logic Control of Drives and Power Systems, Newnes Press, Great Britain.
5) Fuzzy Logic Control Tool Box, User's Guide Version 7.5.0.342 (R2007). August 15, 2007, the Maths Works Inc.
6) Goulias D.G. (2001). "Reinforced recycled polymer based composites for highway poles", Journal of Solid Waste Technology and Management, **27**(2). muse.wider.edu/~sxw0004/cumindex.html
7) Gronostajski, J. and Matuszak, A.A. (1999). —The Recycling of Metals by Plastic Deformation: An Example of Recycling of Aluminum and its Alloys Chips". Journal of Materials Processing Technology. 92:35-41.
8) Kufman, S. M. (2004) Analysis of Technology And Infrastructure f the Paper Recycling Industry in New York City. M.S. Thesis, Department of Earth and Environmental Engineering Fu Foundation of School of Engineering and Applied Science, Columbia University.
9) Kumar, R. (2007), Optimal Blending Model for Paper Manufacturing With Competing Input Materials Pati, POMS 18th Annual Conference, Dallas, Texas, U.S.A.
10) Kumaran, D.S., Ong. S.K., Tan, R.B.H. and Nee, A.Y.C. (2001). "Environmental lifecycle cost analysis of products", Environmental Management and Health, **12**(3), 260-276.
11) Lee C.C. (1995). —Fuzzy Logic in Control System: Fuzzy Logic Controller Part I and Part II". IEEE Trans Systems Man and Cybernetics. 20:404 – 418.
12) Mamdani, E.H. (1977), Application of fuzzy logic to approximate reasoning using linguistic systems. Fuzzy Sets and Systems 26, 1182–1191.
13) MisrawatiMisman, SharifahRafidah Wan Alwi and Zainuddin Abdul Manan. (2008), State-Of-The-Art for Paper Recycling. International Conference on Science and Technology (ICSTIE),
14) UniversitiTeknologi MARA, Pulau Pinang, Malaysia,
15) Obot, O. U. (2008). Fuzzy rule-base frame work for the management of tropical diseases. Int. Journal of Medical Engineering and Informarics, 1(1): 7 – 17.
16) Oke, S.A., A.O. Johnson, I.O. Popoola, O.E. Charles-Owaba, and F.A. Oyawale. (2006). —Application of Fuzzy Logic Concept to Profitability Quantification in Plastic Recycling". Pacific Journal of Science and Technology. 7(2):163-175.
17) Pulp and Paper Information Centre Website, (2008). http//www.ppic.org.uk Accessed on 22 March 2008.
18) Samakovlis, E., (2004), Revaluing the hierarchy of paper recycling. Energy Economics, 26,101-122, .
19) Saritas, I. A. and Sert. U. (2003), A Fuzzy expert system design for diagnosis of prostrate caner. Proceedings of International onference on Computer Systems and Technologies, Kanya, Turkey.
20) TAPPI - The Leading Technical Association for the Worldwide Pulp, Paper and Converting Industry (2001). —How to recycle paper?".
21) Treloar, G.J., Gupta, H., Love, P.E.D. and Nguyen, B., (2003), An analysis of factors influencing waste minimization and use of recycled materials for the construction of residential buildings, International Journal of Management of Environmental Quality, 14(1), 34-145.
22) Wang, J (2005), A Fuzzy Vacuum Cleaner Hardware and Software Design. M.Sc. Thesis, Department of Automatic Control and System Engineering, University of Sheffield.
23) Wongthatsanekorn, W. (2009) A Goal Programming Approach for Plastic Recycling System in Thailand. Proceedings of World Academy of Science, Engineering and Technology Volume 37, Issn 2070-3740
24) Zadeh, L. A. (1965), Fuzzy Sets, Information and Control, 1965
25) http://www.recyclingtoday.com, http://www.recyclinginternational.com,

Future Of Human Security Based On Computational Intelligence Using Palm Vein Technology

GJCST Classification
I.2.m.C.2.0

T. VenkatNarayana Rao[1], K.Preethi[2]

Abstract **-This paper discusses the contact less palm veinauthentication device that uses blood vessel patterns as a personal identifying factor. The vein information is hard to duplicate since veins are internal to the human Body. This paper presents a review on the palm vein authentication process and its relevance and competence as compared to the contemporary Biometric methods. Th is authentication technology offers a high level of Accuracy. The importance of biometrics in the current field of Security has been illustrated in this paper. We have also outlined opinions about the utility of biometric authentication systems, comparison between different techniques and their advantages and disadvantage. Its significance is studied in this paper with reference to the banks, E-Voting, point of sale outlets and card/document less security system. Fujitsu plans to further expand applications for this technology by downsizing the sensor and improving the certification speed. I.2.m C.2.0**

Key words-infrared rays, pattern, contact less ,deoxidized hemoglobin , sensors.

I. INTRODUCTION

The prime responsibility of any technological development is to provide a unique and secure identity for citizens, customers or stake holders and it is a major challenge for public and private sector organizations. The rise of identity theft in the internet age is well documented. Recent figures reported a 40% increase in the number of victims of impersonation during the last one year, when compared with the same period in 2009 . Organizations hold large volumes of personal data and thus entail flawless protection. The pattern of blood veins is unique to every individual human, and same is the case among similar twins also. Palms have a broad and complicated vascular pattern and thus contain plenty of differentiating features for personal identification. It will not vary during the person's lifetime. It is very secure
method of authentication because this blood vein pattern lies underneath human skin. This makes it almost impossible for others to read or copy the vein patterns. An Image pattern of a human is captured (Figure 1) by radiating his/her hand with near-infrared rays. The reflection method illuminates the palm using an infrared ray and captures the light given off by the region after diffusion through the palm. The underlying technology of palm-vein biometrics works by extracting the characteristics of veins in the form of a bit image database [1][4]. As veins are internal in the body and .

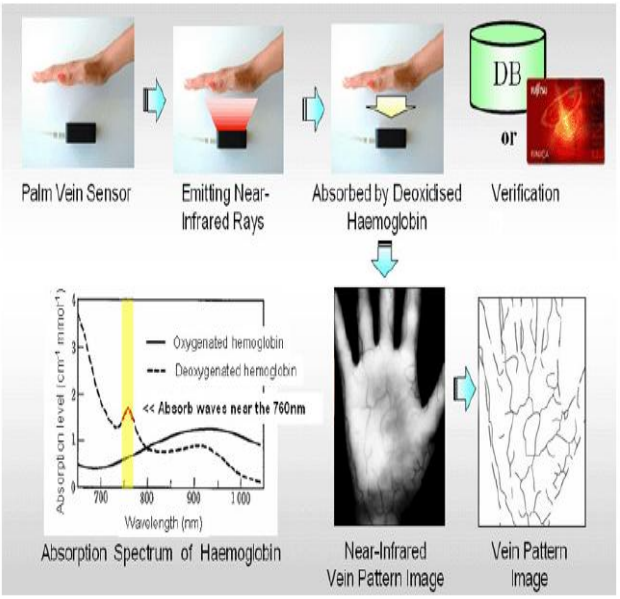


Figure 1 Flow of Palm Vein Technology Process [16].

Biometric template - a numeric representation of several characteristics measured from the captured image, including the proximity and complexity between intervened veins (figure 1). This template is then used to compare against a user's palm scan each time they undergo authentication process. This technology is nonintrusive i.e. the user need not physically touch the sensor. The users must hold their hand above the sensor for a second. The method is also highly accurate. The International Biometrics Group (IBG), which evaluates all types of biometrics products through comparative testing, found that palm-vein technology was on par with iris scan biometrics in accuracy ratings. Palm-vein recognition technology is notably less costly than iris scanning technology. In fact, the only biometric solution less expensive than palm-vein authentication is fingerprint recognition but it has its own overheads on securityfeature. For health care organizations, effective palmveinrecognition solutions enable accurate identification of patients, enabling them to quickly retrieve their electronic medical records when they check into respective hospitals. This eliminates the potential human error of accessing the erroneous record, thus helping in protecting patients from identifying fraudulent attempts . Until now, there has been no biometric

About[1]-Professor and Head, C.S.E, Tirumala Engineering College, Bogaram, A.P, India tvnrbobby@yahoo.com
About[2]- Assistant Professor, C.S.E, Tirumala Engineering College, Bogaram, A.P, India preethi.kona@gmail.com

technology thatcan achieve the highest levels of security and usability at a reasonable cost. Palmvein recognition hits that success spot of biometrics between security, cost, accuracy and ease of use that make it an optimal answer and IT enabled control solution for health care organizations and hospitals. Compared with a finger [4] or the back of a hand, a palm has a broader and more complicated vascular pattern and thus contains a wealth of
differentiating features for personal identification. The palm is an ideal part of the body for this technology; it normally does not have hair which can be an obstacle for photographing the blood vessel pattern, and it is less susceptible to a change in skin color, unlike a finger or the back of a hand. However research appears to have conquered this challenge and an early demonstration device is built into a computer mouse by Fujitsu in a development of vein pattern identification by researcher Masaki Watanabe. This was used to control access to the computer system. More recently, Fujitsu demonstrated their Contact less Palm Vein Identification System at the annual CeBIT show in March 2005. At least five vendors have been pursuing this technology including Fujitsu, Hitachi, Bionics Co., Identica and Techsphere. Japan‘s Bank of Tokyo-Mitsubishi made this technology available to customers on 5000 ATM‘s from October 2004. The biometric template is stored on a multi-purpose smart card that also functions as a credit and debit card and issued to customers. Other Japanese banks are also now introducing this technology. EFTPOS terminals, incorporating palm vein technology are being developed for use in for use in retail stores. While the size of earlier devices limited their use and added to cost, recent developments have reduced the size to make mobile and portable devices feasible. These use 35mm sensors which makes the device small enough to use with laptops and other mobile devices and other office equipment such as copiers [8]. Several of Japan's major banks have been using palm and finger vein recognition at cash points, rather than PIN, for almost 3 years now and are confirming extraordinarily high standards of accuracy.

II. PRINCIPLES OF PALM VEIN BIOMETRICS AND CONTACT LESS AUTHENTICATION

The contact less palm vein authentication technologconsists of image sensing and software technology. Thepalm vein sensor (Fig.2) captures an infrared ray image of the user‘s palm. The lighting of the infrared ray is controlled depending on the illumination around the sensor, and the sensor is able to capture the palm image regardless of the position and movement of the palm. The software then matches the translated vein pattern with the registered pattern, while measuring the position and orientation of the palm by a pattern matching method. In addition, sufficient consideration was given to individuals who are reluctant to come into direct contact with publicly used devices [7] [14]. The deoxidized hemoglobin in the vein vessels absorbs light having a wavelength of about 7.6 x 10-4 mm within the nearinfraredarea. The device captures an image of veinpatterns in wrist, palm, back of the hand, finger or face. This is similar to the technique used to capture retinal patterns. The backs of hands and palms have more complex vascular patterns than fingers and provide more distinct features for pattern matching and authentication. As with other biometric identification approaches, vein patterns are considered to be time invariant and sufficiently distinct to clearly identify an individual. The difficulty is that veins move and flex as blood is pumped around the human body[12]. Human Physiological and behavioral characteristic can be used as a biometric characteristic as long as it satisfies the following requirements:

- Universality: each person should have the characteristic.
- Distinctiveness: any two persons should be sufficiently different interms of the characteristic.
- Permanence: the characteristic should be sufficiently invariant (with respect to the matching criterion) over a period of time.
  Collectability: the characteristic can be
- measured quantitatively.

How does Biometrics System Work?

Irrespective of type of biometric scheme is used; all have to go through the same process. The steps of the process are capture, process, and comparison.

• Capture – A biometric scheme is used to capture a behavioral or physiological feature.

• Process – The captured feature is then processed to extract the unique element(s) that corresponds to that certain person

• Comparison – The individual is then enrolled into a system as an authorized user. During this step of the process, the image captured is checked against existing unique elements. This verifies that the element is a newly authorized user. Once everything is done, the element can be used for future comparisons [5].

Certain questions need to be asked when choosing a Biometric System Implementation:

1) What is the level of security is needed?
2) Will the system be attended or unattended?
3) Does your requirement demand resistance to spoofing?
4) What reliability level is required?
5) Should this system be made available through outtheday?
6) Does the system require backups- if yes how many hours of Backup?
7) What is the acceptable time for enrollment?
8) Is privacy to be addressed for your system?
9) What about the storage of the signature?
10) Is the system integrated with Front end and Backenddatabase system?
11) Is the system open for Maintenance activity and tuning around the clock?

In practice, a sensor emits these rays and captures an image based on the reflection from the palm. As the hemoglobin absorbs the rays, it creates a distortion in the reflection light so the sensor can capture an image that accurately records the unique vein patterns in a person's hand. The recorded image is then converted to a mathematically manipulative representation of bits which is highly complicated to get

forged or compromised. Based on this feature, the vein authentication device translates the black lines of the infrared ray image as the blood vessel pattern of the palm (Figure 2), and then matches it with the previously registered blood vessel pattern of the individual [9].

1) *Biometrics parameters and keywords of palm vein technology*

• **Vein patterns:** Distinctive and unique to individuals, Difficult to forge
• **False acceptance rate:** A rate at which some one other than the actual person is recognized
• **False rejection rate: A** rate at which the actual person is not recognized accurately
• **Potential is limitless:** Easy to install on personal computer, Reliable , Accurate, Fast, Small
• **Equal Error Rate (EER):** Point where FAR=FRR
• **Failure to Enroll Rate (FTER):** Percentage of failures to enroll of the total number of enrollment attempts.

III. THE WORKING MECHANISM/ IMPLEMENTATIO BEHIND PALM VEIN BIOMETRIC

An individual's palm vein image is converted byalgorithms into data points, which is then compressed, encrypted, and stored by the software and registered long with the other details in his profile as a reference for future comparison (figure 2). Then, each time a person logs in attempting to gain access by a palm scan to a particular bank account or secured entryway, etc., the newly captured image is likewise processed and compared to the registered one or to the bank of stored files for verification, all in a period of seconds. Implementation of a contact less identification system enables applications in public places or in environments where hygiene standards are required, such as in medical applications. The vein pattern is then verified against a reregistered pattern to authenticate the individual. Numbers and positions of veins and their crossing points are all compared and, depending on verification, the person is either granted or denied access. As veins are internal in the body and have a wealth of differentiating features, attempts to forge an identity are extremely difficult, thereby enabling a high level of security [10]. In addition, the sensor of the palm vein device can only recognize the pattern if the deoxidized hemoglobin is traverse through the veins of the hand which makes the process more secured and safe.

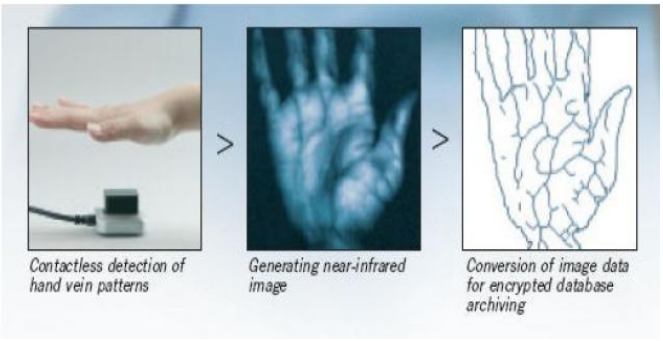


Figure 2. Palm Exposure to Sensor

And Conversion/Comparison against from Archival Database

1) *Advantages and Disadvantages of Palm vein technology*

ADVANTAGES	DISADVANTAGES
It does not require user contact	----------
Matching performance is high	----------
Most suitable for authentication	----------
It is accurate , Potential is limitless	Require specialized devices, so can be expensive as of now.
Easy to use or handle	Requires highly active deoxidized hemoglobin.
Unlike fingerprints that change during childhood, the palm vein pattern is established in the womb and is constant throughout a person's life.	---------
It is neither be stolen nor reproduced.	----------

Table: 1 - Advantages and Disadvantages of Palm Vein Technology

IV. PRACTICAL APPLICATIONS OF PALM VEIN BIOMETRICS

The rapid growth in the use of e-commerce and onlineapplications requires reliable user identification for effective and secure access control. Palm vein identification has emerged as a promising component of biometrics study. .Applications of palm vein biometrics are: Security systems, Log-in control or network access, Healthcare and medical record verification , electronic record management; Banking and financial services like access to ATM , kiosks, vault etc. The medical problems like diabetes, hypertension, atherosclerosis, metabolic disorders and tumors are som e diseases which affect the vascular systems and are need to be attended very often by the doctor and palm vein technology can come as a bonus facility for faster and accurate medicalreading.In this following section, we present a briefreview on the applications and features of applications of palm vein technology useful in the above mentioned sectors.

1) *Palm Vein for Financial SecuritySolutions*

A rapidly increasing problem among financial sectors in Japan is the illegal withdrawal of bank funds using stolen or skimmed fake bankcards. To address this, palm vein authentication has been utilized for customer confirmation of transactions at bank windows or ATMs.The smart card from the customer‘s bank account contains the customer‘s palm vein pattern and the matching software of the palm vein patterns. A palm vein authentication device at the ATM

(Figure 3) scans the customer‘s palm vein pattern and transfers it into the smart card. The customer‘s palm vein pattern is then matched with the registered vein pattern in the smart card. Since the registered customer‘s palm vein pattern is not released from the smart card, the security of the customer‘s vein pattern is preserved. In 2004, the Suruga Bank and the Bank of Tokyo-Mitsubishi in Japan deployed a secured account service utilizing the contactless palm vein authentication system. Several other banks in Japan have followed suit in 2005[13][17]. Fujitsu plans to develop another type of ATM (Figure 3) for use at convenience stores in Japan, embedding the palm vein authentication sensor in the ATM.

Figure 3. ATM with palm vein access control unit pattern authentication sensor unit

2) *Access control in house hold andBusiness Houses*

The palm vein pattern sensor is also used for access control units. The ―palm vein authentication access control device‖ is comprised of the palm vein pattern sensor, a keypad and a small display. This device controls access to rooms or buildings that are for restricted personnel. The device consists of two parts: the palm vein sensor, plus the control unit that executes the authentication processing and sends the unlock instruction [15]. A simple configuration system can be achieved by connecting this device to the electric lock control board or electric locks provided by the manufacturer.

3) *E-Voting*

The physical traits of an individual confirm or verify theiridentity. This gives rise to ensure citizens e-Voting to brfool proof with no flaws, thus can be employed widely for unique security benefits for identification and security. They can reduce and in some cases eliminate the need for individuals to carry documentation or other physical security measures they might lose or to remember passwords to prove their identification. A more secure future: enabling security through biometrics. Palm vein technology can be a good alternative to world in federal and general election system to figure out undisputed mandate to a winning party. This can introduce much accuracy and reliability dealing millions of voters with inhours unlike classical manual methods of franchise votes.

4) *Nations Border Security Control*

Any Border officers have traditional methods by comparing an individual‘s passport photo to the person.in front of them. Many supporting documents such as entry visas carry no identification other than names, passport numbers, date of birth and addresses etc. Introduction of Biometrics can bring about revolutionary changes in eliminating intrusion into nation‘s entry. The palm vein technology along with face recognition and fingerprint biometrics can ease identifying fraudulent and terrorist groups from creeping into other countries.

5) *Retail Industry*

Big retail outlets are making use of biometrics to cater to huge flock of customers and timely delivery of its products and services. This can regulate children age on the purchase of restricted product such as pharmaceuticals, digital products such as alcohol and tobacco etc. If Biometrics is employed in industries along with the ERP systems it can directly address and minimize the commercial and public sector security check burden for dispensing services its products. This can reduce the role of huge server records retrieval and verification at source.

V. RECENT TECHNOLOGICAL DEVELOPMENTS USING PALM VEIN BIOMETRIC AUTHENTICATION SENSORS

Fujitsu Limited and Fujitsu Frontech Limited [17], Japan has announced that they have developed a PC Login Kitfor use with the Palm Secure palm vein biometric authentication device and begun sales of a mouse model and a standard model for corporate users. Palm Secure PC Login Kit comes standard with loginauthenticationsoftware, enabling client-side authentication and eliminating the need to use an authentication server, which had been required up until now [11]. In addition, other improvements have been incorporated, such as faster authentication speeds without a palm guide and greater tolerance for the distance and angle of the handwhen it passes over the device. With the new PalmSecure PC Login Kit, logins to PCs or applications that are in use until now required IDs and passwords can now be done using the highly secure palm vein biometricauthentication method. In recent years, as part of efforts to comply with Japan‘s Personal Information Protection Law and enhanced internal corporate compliance policies, it has become increasingly important to authenticate the identity of people using particular PCs in order to prevent data leaks from PCs that occur because of unauthorized access or identity fraud. Since 2004, Fujitsu and Fujitsu [17] Frontech commercialized the Palm Secure palm vein biometric authentication device, which offers superior security and is easy to use. Since then, the

companies have provided the technology to financial institutions and wide array of other industries and organizations for use in various applications,
including login to PCs, physical admission into secured areas, management for work time clocks, and library book lending systems. The two companies developed Palm Secure PC Login Kit to make it more simple and economical for customers to deploy Fujitsu's sophisticated palm vein authentication technology. Installing loginauthentication software as standardequippedsoftware, sophisticated authentication can be handled by the PC itself, with no need for an authentication server.Palm secure is now widely used in various fields: ATM, 92% of all Japanese ATMs i.e. 18,000 + ATM machines for Bank of Tokyo – Mitsubishi. The mouse model, which is the world's first PC mouse equipped with a palm vein biometric authentication sensor, can easily replace an existing PC mouse, offering convenience and space-saving advantages. The companies have also added a compact and portable standard model to their line of PC login kits for house hold security , user identification and passport verification systems. Both the mouse and standard models are available in black, white and gray to coordinate with different offices and computers. Fujitsu Frontech is in charge of development and manufacturing of the PalmSecure PC Login Kit, with both Fujitsu and Fujitsu Frontech handling sales. Over the next three years, Fujitsu aims to sell 200,000 PalmSecure sensors of all types globally [12][17].

VI. Result of Experiments

As a result of the Fujitsu research using data from 140,000 palms (70,000 individuals), Fujitsu has confirmed that the FAR is 0.00008% and the FRR is 0.01%, with the following condition: a person must hold the palm over the sensor for three scans during registration, and then only one final scan is permitted to confirm authentication. In addition, the following data has been used to confirm the accuracy of this technology: data from 5-year to 85-year old people of various backgrounds based on statistics from the Ministry of Internal Affairs and Communications of Japan's population distribution; data from foreigners in Japan based on the world population Distribution announced by the U.N.; data of the daily changes of Fujitsu employees tracked over several years; and Data of various human activities such as drinking, bathing, going outside, and waking up. Figure 4 showcases the acceptance and rejection FRR (False Acceptance Rate) and FAR (FalsRejection Rate) criteria's mapped with the error rate permissible. Its is very much evident from the table Table2 how secure and efficient is Palm vein technology over other technologies.

TECHNOLOGY	FALSE ACCEPTANCE RATE	FALSE REJECTION RATE
Palm Secure	.00008%	.01%
Fingerprint	1-2%	3%
Iris	.0001% - .94%	.99% - .2%
Voice	2%	10%

Table :2. Comparison of various Biometric Technologies w.r.t FRR and FAR.

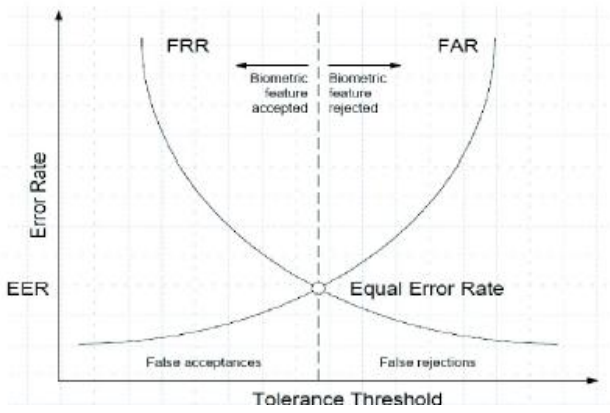


Figure 4. Performance Evaluation

VII. Conclusion

Applications of palm vein biometrics are: a. Security systems: physical admission into secured areas; b.Log-in control: network or PC access; c. Healthcare: ID verification for medical equipment, electronic record management; d. banking and financial services: access to ATM, kiosks, vault. We have already started the work which can be useful for any one of the above mentioned sectors. Biometrics is used for identification purposes and are usually classified as physiological or behavioral.Sometimes a certain biometric can be classified as both. As we continue to progress into the future, more and more biometric schemes will become available. Also, more of the existing biometric schemes will advance further for a higher level of security. Identification and verification classify biometrics even further. The identification process matches 1 to N and the verification process is 1 to 1. As the need for security increases, so will the need for biometrics. It will definitely be interesting to see what the future holds for palm vein biometrics. Palm Vein Technology has presented a new face to the world of security system. It has low FAR and FRR and it has emerged as more hygienic as compared to other systems. In future it can be combined with multimodal biometric system to make the system more attack proof. Thus, we can look forward for an extra ordinary biometric based security systems which would include evenpasswords along with watermarking authentication algorithms.

VIII. REFERENCES

1) S.-K. Im, H.-M. Park, S.-W. Kim, C.-K. Chung, and H.-S.Choi, —Improved vein pattern extracting algorithm and itsimplementation," Proc. Int. Conf. Consumer Electronics, pp. 2-3, Jun. 2000.
2) S. K. Im, H. M. Park, Y.W. Kim, S. C. Han, S.W. Kim, and C.Hang, —A biometric identification system by extracting hanvein patterns," J. Korean Phys. Soc., vol. 38, pp. 268–272, Mar.2001.
3) T. Tanaka and N. Kubo, —Biometric authentication by handvein patterns," Proc. SICE Annual Conference, Yokohama,Japan, pp. 249-253, Aug. 2004.
4) G. T. Park, S. K. Im, and H. S. Choi, —A person identificationalgorithm utilizing hand vein pattern," Proc. Korea SignalProcessing Conference, vol. 10, no. 1, pp. 1107-1110, 1997. 24
5) S. Zhao, Y. Wang and Y. Wang, —Biometric verification byextracting hand vein patterns from low-quality images,"Proc.4th Intl. Conf. ICIG, pp. 667-671, Aug. 2007.
6) http://www.viewse.com.cn/ProductOne.asp?ID=10 6 13. Y.Ding, D. Zhuang and K. Wang, —A study of hand veinrecognition method," Proc. IEEE Intl. Conf. Mechatronics &Automation, Niagara Falls, Canada, pp. 2106 – 2110,Jul. 2005.
7) K. Wang, Y. Zhang, Z. Yuan, and D. Zhuang, —Hand veinrecognition based on multi supplemental features of multiclassifier fusion decision," Proc. IEEE Intl. Conf. Mechatronics& Automation, Luoyang, China, pp. 1790 –1795, June. 2006.
8) L. Wang and G. Leedham, —Near and Far-Infrared imagingfor vein pattern biometrics," Proc. IEEE Intl conf. Video& Signalbased Surveillance, AVSS'06, Sydney, pp. 52-57, Nov. 2006.
9) Handbook of Biometrics, A. K. Jain, P. Flynn, and A. Ross(Eds), Springer, 2007.
10) L. Wang, G. Leedham and Siu-Yeung Cho, —Minutiae Feature Analysis for Infrared Hand Vein Pattern Biometrics," Pattern Recognition, 41 (3), pp. 920-929, 2008.
11) J.-G. Wang, W.-Y. Yau, A. Suwandy and E. Sung, —Person recognition by palmprint and palm vein images based on 'Laplacianpalm' representation," Pattern Recognition, vol. 41,pp. 1531-1544, 2008.
12) http://www.fujitsu.com/global/about/rd/200506palmsvein.html
13) Ding, 05 Yuhang Ding, Dayan Zhuang and KejunWang,—A Study of Hand Vein Recognition Method", The IEEEInternational Conference on Mechatronics & AutomationNiagara Falls, Canada, July 2005, pp. 2106-2110.
14) Tanaka, 04 Toshiyuki Tanaka, Naohiko Kubo, —BiometricAuthentication by Hand Vein Patterns", SICE AnnualConference, Sapporo, August 4-6, 2004, pp. 249-253.
15) Bio-informatics Visualization Technology committee, BioinformaticsVisualization Technology (Corona \Publishing, 1997), p.83, Fig.3.2.
16) www.palmsure.com/technology.
17) —Fujitsu Palm Vein Technology," Fujitsu, May 2005,Available at http://www.fujitsu/globalabout/rd/200506palmvein html.

Digital Literacy: The Criteria For Being EducatedIn Information Society

GJCST Classification
J.1

Abdul Sattar Khan, Allah Nawaz , shadiullah and qamarafaq

***Abstract*-Digital (computer) literacy is the new title for „educated". Both teachers and students have no option but to acquire a level of computer-literacy to catch up with the growing digital societies. Governments and higher education institutions (HEIs) are making all out efforts by providing eLearning environments to gain some levels of digital literacy of the masses at large and the university-constituents. Both developed and developing states are trying to figure out a required digital literacy curriculum for the training of teachers and the students. However, given that there are several meanings of computer-literacy therefore; research is going on about the contents of the curriculum and the pedagogical requirements of ICT-education. Furthermore, the concepts of global-village, globalization, information or knowledge society, ePedagogy, eStudents and eCourses – all are casting increasing pressures on the academicians, HEIs and governments to take digitial opportunity initiatives (DOI) for digital-literacy of the masses to generate workforce for the eGovernment, eCommerce and eLearning.Research reveals that learners hold different perceptions about the nature and role of ICTs such as: instrumental and substantive. Some consider it just like any other technology with no value-implications for the learner and society. Substantive theorists however believe in the determinist role of technologies for changing the society. Whatever the paradigm, learners are facing several hurdles in acquiring digital command like perceptual differences, demographic diversities, resistance to change, training issues, and so on. However, most of the researchers are coming up with the findings that, perceptions, theories, teaching/learning styles of the teachers, students and other stakeholders play decisive and determinist role in determining the speed and quality of computer-literacy.It is well-documented that the contents and dynamics of computer-literacy in any state depend on the objectives to be realized through ICTs. Depending on the perceptions about eLearning, technologies are either used to achieve immediate objectives for instant contributions (instrumental-view) or long-term and broader objectives (substantive or liberal-view). It is argued that none of the instrumental or substantive views are good or bad rather two stages or steps in the evolution of eLearning from objectivist thinking to social constructivist digital platforms. Almost every country and HEI is first experimenting with the instrumental benefits of ICTs and this practice is more rampant in the developing countries. This paper is an effort to draw a picturesque of digital-literacy in the background of HEIs.**

Keywords-Digital/Computer-Literacy,Educational Technologies, Paradigm, Instrumental, Substantive, Objectivist, Cognitive and Social Constructivist, ePedagogy, eStudent, eCourse, DOI, HEI,

About- assitant professor, Institution/University:gomal university deraismaeel khan ..KPK sattarmpadik@gmail.com

I. Introduction

The universal demand for _computer-literacy' emanates from the dominance of ICTs in different aspects of contemporary life (Oliver, 2002). The supporters of _social inclusion through ICTs', emphasize _electronic-literacy' as a key to bridge digital-divide (Macleod, 2005). Digital literacy is deemed necessary for —mindful learning in the information society (Aviram&Eshet-Alkalai, 2006)." Students, teachers, and employees define computer literacy differently (Johnson et al., 2006) however, commonly; people acquire their _technology-literacy' either formally through formal courses or informally at home, from friends, or by themselves (Ezziane, 2007).The indispensability of digital literacy is evident from the findings and arguments of researchers around the globe. For example, ICTs (connectivity-tools) have been found helpful in reducing the problems of _isolation' (Tinio, 2002; Abrami et al., 2006; Vrana, 2007) and _disempowerment' (Macleod, 2005; Wims& Lawler, 2007) for the developing countries and marginalized groups. Digital opportunity initiatives (DOI) are proving powerful tools for _poverty-alleviation' and _economic-development' in developing states (Macleod, 2005; Hameed, 2007; HEC, 2008). Developing countries like Pakistan are entering into _international and national' partnerships to capitalize on global ICT-resources (Tinio, 2002; Mathur, 2006; Baumeister, 2006; Kopyc, 2007). Furthermore, within university environment, eLearning tools create _learner-centric' and _collaborative-learning environments' where they are empowered to self-control their learning processes (Mejias, 2006). The expectations of employers, parents, and educators from the graduates (about digital literacy) are changing (Johnson et al., 2006). Therefore, most of the universities have started compulsory computer literacy courses however, to provide required command over computers, it is important to determine a _customized digital curricula and ePedagpgy' (Martin &Dunsworth, 2007). However, in third world countries, very little research has been published about students' perceptions of their computer literacy (Bataineh& Abdel-Rahman, 2006).Thus, digital literacy is not only shifting power bases in the developing countries from —elites to masses (Macleod, 2005)" rather it is increasingly —perceived as a survival skill (Aviram&Eshet-Alkalai, 2006)." However, acquisition of computer-literacy knowledge and skills is neither automatic nor simple rather dependent on a variety of personal (teacher, students, administrators), organizational (higher education institution – HEI) and broader political and social factors (local, national and international) within which eLearning occurs. Following analysis and discussion unfolds the concept, learning

paradigms and barriers in digitizing the communities inhabiting modern _information and knowledge societies'.

II. DIGITAL LITERACY

The illiterate of the 21st century are not those who cannot read and write, but those who cannot learn, unlearn, and relearn (Tinio, 2002). The definition of computer literacy has evolved overtime as technology improved and society became more dependent on computers. Some 50 years ago when a computer nearly filled a room, computer literacy meant being able to program a computer (Johnson et al., 2006). Today, when every user holds a computer, computer literacy is defined as an understanding of computer characteristics, capabilities, and applications, as well as an ability to implement this knowledge in the skillful, productive use of computers in a personalized manner (Martin &Dunsworth, 2007). Terms such as computer competency, computer proficiency, and computer literacy are used interchangeably (Johnson et al., 2006).With today's technological society, basic computer literacy is emphasized in every institution (Ezziane, 2007). Digital literacy is a combination of technical-procedural, cognitive and emotional-social skills, for example, using a computer involves procedural skills (file-management), cognitive skills (intuitively reading the visual messages in graphic user interfaces) (Aviram&Eshet-Alkalai, 2006). With the changes in technology, the elements of computer literacy are constantly changing and thus, educators must constantly revise the course to include the latest technological trends (Martin &Dunsworth, 2007).

1) *Elearning*

eLearning is widely researched in the perspectives of –higher education as well as corporate training (Tinio, 2002)" and explained as the 'application of electronic technologies' in enhancing and delivering education (Gray et al., 2003). ICTs represent computers, networks, software, Internet, wireless and mobile technologies to access, analyze, create, distribute, exchange and use facts and figures in a manner that has been unimaginable hitherto (Beebe, 2004). A variety of concepts is interchangeably used to represent eLearning including: computer-based instruction, computer-assisted instruction, web-based learning, electronic learning, distance education, online instruction, multimedia instruction, and networked learning are a few (Tinio, 2002; Abrami et al., 2006; Baumeister, 2006; Manochehr, 2007; Sife et al., 2007; Wikipedia, 2009). In eLearning the data-networks such as, internet, intranet, and extranet are used to deliver course contents and facilitate teachers, students and administrators (Tinio, 2002). The term networked learning is also used as a synonym for eLearning (Baumeister, 2006). Internet and web-based applications are most widely used educational technologies in the eLearning systems (Luck & Norton, 2005) therefore; teachers, students and education managers are using the web for a variety of purposes (Manochehr, 2007).The concept of eLearning also has non-educational conceptions. Hans-Peter Baumeister (2006) notes that the meaning of eLearning varies with a change in the context: Political dimension denotes the modernization of whole education system; but Economic view defines eLearning as a sector of eBusiness. In nutshell, eLearning begins with a partial or supplementary use of ICTs in classroom then steps into a blended or hybrid use and finally offers online synchronous and asynchronous virtual learning environments serving physically dispersed learners (Sife et al., 2007).

2) *Educational Technologies*

ICTs refer not only to modern hi-tech computers and networks rather there are old and new ICTs where radio, television, telephone, fax, telegram, etc are now old while computer-networks, Internet, e-mail, and mobile learning are new tools (Hameed, 2007). At the same time, eLearning technologies are burgeoning in terms of hardware, software and a variety of applications in education for teachers, students and administrators. Educational technologies come in variety (Sife et al., 2007) however, computers, networking and hypermedia is the core paradigms for different roles of eLearning (Ezziane, 2007).

- ***Computer***

The primary tool for eLearning is the computer, which has traveled a long way since 1960s when UNIVAC in USA and Baby-Computer in UK emerged as the pioneers of a technology, which is now controlling almost every aspect of human life. The transformation from XT (extended-technology) to AT (advanced-technology) or Personal Computer (PC) in 1980 was the second big innovation making computers _a personal gadget' for everybody and anybody. A computer is an intelligent-machine and a powerhouse for users in terms of its processing capabilities and speed (i.e., user command is executed on a click), storage capacity (hard-disk and from floppy to flash and XDrives), and graphic interfaces (i.e., graphical-user-interface GUI) to interact with different parts of the machine, like, activating a software, using CD-drive, printing a document or picture, copying a file from hard disk on a _data-traveler.'

- ***Networking***

When computers are wired together for communication and resource-sharing, it is called a digital network. Networking has elevated the role of ICTs and a huge body of research is underway to make connectivity more and more powerful. Networking is evolving from simple networks into complicated forms of Internet, intranet and extranet along with web-technologies thereby converting the world into a _global-village'. Networking eliminates the geographical and physical constraints through a multitude of tools and techniques based on the communication-protocol of TCP/IP, onto which Internet is anchored. According to Glogoff (2005) a network is a platform (internet, intranets and extranets) decorated with web-based tools of hypermedia and multimedia applications managed through learning and content management systems (LMS, LCMS). It is therefore evident that Internet is becoming an indispensable tool for learning and social life (Barnes et al., 2007).It is reported that that many of the eLearning facilities in HEIs offer traditional print syllabus via Internet however many researchers assert that innovative applications of Web are

diverse (Wood, 2004). Likewise, John Thompson (2007) notes that accessing the Internet is like going to the library for a book however, Internet offers opportunities which need to be explored the technologies are designed well and used as intended (Wijekumar, 2005). Internet technologies (now offering Web 2.0, such as blogs, wikis, RSS, podcasting etc.), virtual reality applications, videogames and mobile devices are some of the many innovations, which are common in daily life for communication and entertainment are equally helpful in learning and emerging as such (Chan & Lee, 2007). Through Web 2.0 technologies, users can communicate and interact globally via internet in a paradigm of open communication, decentralization of authority and freedom to share and re-use online resources (Wikipedia, 2009).

3) *Curricula for Digital Literacy*

The ‗curricula' of any country is viewed as ―a snapshot of the current state of knowledge (Ezer, 2006).‖ Therefore, the debate about whether education should be focused on the current job market (instrumental) or intellectual attainment (liberal) is ongoing. It is reported that most of the current computer-training and education is ineffective because it is more technical and less concerned with the contexts and real world problems (Ezer, 2006). Due to increased demand for ICT-professionals, the universities across the world have responded by developing programs without ―an existing model for guidance (Ekstrom et al., 2006).‖ However, Stephen J. Andriole (2006) warns that ―the gap between what we teach and what we do is widening … academic programs should acknowledge the widening gap between theory and practice, especially since it has enormous implications for their graduates' ability to find work.‖Despite some similarities in the computing curricula there are clear distinctions of being developed and developing countries. In a comparative study of the computing curricula in India and America, the researcher found that there are similarities in terms of offering fundamental courses in IT, system development, basics of operating systems, hardware architecture, web technologies and programming fundamentals. However, the differences are more obvious for example; India is more instrumental while American education is more liberal in computing curricula with less emphasis on hard sciences than Indian curriculum (Ezer, 2006).

III. Paradigms for Digital Literacy

It has been found that the use of ICTs is dependant on the perceptions of developers and users about the nature of technologies and their role in different walks of life (Aviram& Tami, 2004). BastienSasseville (2004) have found that ICT-related changes are ―not perceived as a collective experience or social change rather, personal challenge.‖ The literature analysis suggests that two broader theories are discussed over and over saying that ICTs can either play ‗instrumental' or ‗substantive' role in the learning process (Macleod, 2005). Jonathan Ezer, (2006) classifies this issue into ‗instrumental' and ‗liberal' conceptions of eLearning. Instrumental view asserts that ICTs are just technologies and their role depends on their use while substantive view posits that these technologies have the power to change the society and their mere existence can make the difference (Mehra&Mital, 2007).

Three roles of ICTs and digital literacy are suggested (Tinio, 2002):

- Learning *about* ICTs, where digital literacy is the end goal;
- Learning *with* ICTs where technologies facilitates learning; and
- Learning *through* technologies thereby integrating it into curriculum.

Another researcher (Sahay, 2004) identifies four dimensions of computer literacy:

- ICTs as an Object: Learning about the technology itself. Courses are offered to get knowledge and develop skills about different tools. This prepares students for the use of ICTs in education, future occupation and social life.
- Assisting tool: ICT is used as a tool for learning, for example, preparing lectures or assignments, collecting data and documentation, communicating and conducting research. ICTs are applied independently from the subject matter.
- Medium for teaching and learning: This refers to ICT as a tool for teaching and learning itself, the medium through which teachers can teach and learners can learn. Technology based instructional delivery appears in many different forms, such as drill and practice exercises, in simulations and educational networks.
- ICTs for Education Management: The most common and wider application of ICTs is in the organizational and logistic functions of the higher education institutions in the form of transaction processing systems (TPS) and management information systems (MIS).

Given these scenarios, ICTs are either simply a tool (neutral) like any other technology or more than a tool, which can change the people way of life by transforming the education culture (Young, 2003). Research however, reports that ICTs have the potential and flexibility to be used in either ways but as the ICTs become increasingly available to the masses (like internet accessibility) the ICTs begin to affect beyond technical impacts of a tool (Aviram& Tami, 2004). For example, *daily* ‗checking email' has become a common norm even in developing countries. The departure from ‗stand-alone' use of computers to ‗network' applications have increased access to so far inaccessible data sources thereby changing the ‗user-expectations' and thus attitudes to ‗learning-process' itself (Ezziane, 2007).From paradigmatic point of view instrumental vs. substantive reflect the ‗behaviorist vs. constructivist' (Boundourides, 2003) modes of teaching and learning. Behavioral or objectivist approach (instrumental) to teaching and learning ICTs believes more in physical activities and outcomes with the assumption that ‗use makes anything important or otherwise' (Macleod, 2005). On the other extreme,

constructivist (substantive) mode of teaching and learning is ideological and cultural with the belief and conviction that ICTs should be integrated into the very core of teaching and learning with mega changes in pedagogy and knowledge-acquisition (Mehra&Mital, 2007). The technological advancements in eLearning is linked with the theories of learning like behaviorism, objectivism, constructivism, and cognitive and social constructivism (Wikipedia, 2009)."

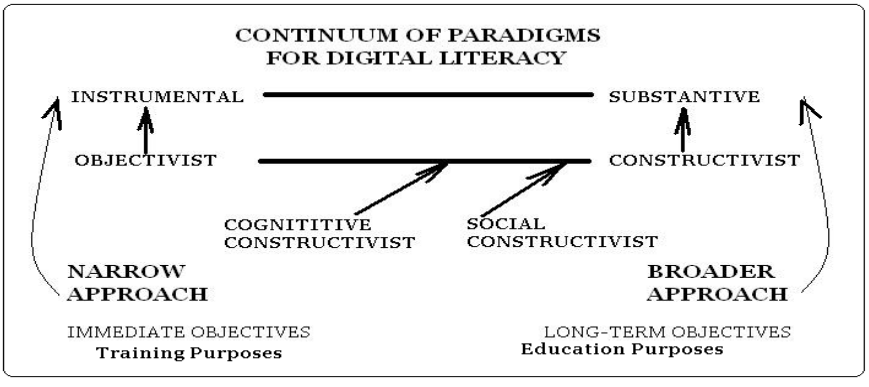
1) *Instrumental/Behaviorist*

Instrumental view of technology is the most commonly held belief, which considers technology as a _tool' without any inherent value (neutral) and its impact lies in how is it used so a _one-size-fits-all' policy of universal employment is used (Macleod, 2005; Radosevich& Kahn, 2006). Instrumental education is based on the premise that education *serves* society so focus is on the utility and usefulness of education to the economy. The underlying philosophy behind the instrumental point of view is the objectivist approach wherein instructor presents the learner with the required stimuli along with the required behavioral responses within an effective reinforcement regime. The degree of learning is assessed through observable measures such as tests, assignments and examinations (Ward et al., 2006)."Objectivism believes that everything related to learning is predictable therefore one learning-model fits all. Likewise, behaviorism give priority to the stimulus-response relationship in learning and underplays cognitive role therefore sees the learning environment as in objectivism (Young, 2003). This is exactly like behavior of scientific management where worker is taken as a part of a big machine called organization. The objectivist teaching gives complete control of materials to the teacher who manages the pace and direction of learning thereby making learning a sequential process where there is a single reality about which the —darners display an understanding through declarative, procedural and conditional knowledge (Phillips et al., 2008)."

2) *Substantive/Constructivist*

The ICTs can play a supplemental as well as central role in learning by providing digital cognitive or adaptive tools or systems to support constructivist learning (Sirkemaa, 2001). Contrary to instrumental, substantive view of ICTs is a determinist or autonomous approach, which argues that technology is not neutral and has positive or negative impacts. Technological determinism encourages the idea that: the mere presence of technology leads to familiar and standard applications of that technology, which in turn bring about social change (Macleod, 2005; Radosevich& Kahn, 2006). The substantive theory matches with the _liberal theory' of education (Ezer, 2006), which views learning as active and interconnected experience and not simply a recollection of facts. This paradigm suggests using ICTs beyond their _supplemental (instrumental)' role to braoder. Constructivists contend that ICTs should not be guided by a technologically deterministic approach rather in the context of social, cultural, political and economic dimensions of using technology so that by facilitating the development of electronic literacy, culturally relevant online content and interfaces and multimedia, the process of social inclusion can be achieved within developing countries (Macleod, 2005). The effectiveness of the behavioral approach is questionable in areas that require comprehension, creativity and 'gray' answers (Ward et al., 2006). The moves towards constructivism in higher education have been pushed by the emergence of universal connectivity through ICTs (Wims& Lawler, 2007), which enabled the masses to globally communicate and most importantly access to the world knowledge resources through the advent of internet after 1990s. Given the access to broader sources of knowledge, contemporary theory suggests that collaborative learning is the most effective means of facilitating teaching and learning in digital environments (Phillips et al., 2008). Furthermore, a new version of this kind of thinking is _social constructivism', which is gaining foothold in higher education because teaching and learning can now easily be undertaken as a social and community activity through social software (Bondarouk, 2006). Social software enables collective learning (social) along with individual (cognitive) with the help of traditional email/chatting and modern wikis, blogs, vblogs, RSS feeds and the list continues(Klamma et al., (2007). For example, RSS is a format used to publish frequently updated works like blog-entries, new headlines, audio and video (Wikipedia, 2009).

Figure 1 Continuum of Paradigms for Digital Literacy



IV. Barriers to Getting Digitally Literate

Given the differences of perceptions (Young, 2003) users behave differently to eLearning tools and techniques for teaching and learning purposes. A key challenge for institutions is overcoming the cultural mindset whereby departments and individuals act as silos, keeping information and control to themselves (LaCour, 2005). Moreover, the training that educators do receive does not always match with their educational needs, because the faculty is rarely involved in the decisions about technology and design of new strategies for technology-integration (Juniu, 2005). In developing countries, ―ICTs have not permeated to a great extent in many higher learning institutions in most developing countries due to many socio-economic and technological circumstances (Sife et al., 2007).‖The greatest challenge in learning environments is to adapt the computer-based system to differently skilled learners. If the environment is too complex the user will be lost, confused or frustrated. On the other hand, too simple or non-systematic environments cause motivational problems (Sirkemaa, 2001). Technology is by nature disruptive, and so, demands new investments of time, money, space, and skills and changes in the way people do things (Aaron et al., 2004). Furthermore, face-to-face communication is critical for classroom social relationships and interpersonal processes while, online technologies have reduced support for social interaction. Although emotions can be conveyed through e-mail or chatting, it does not replace ―the fundamentals of our socio-emotional well-being (Russell, 2005).‖ Thus, ―barriers can make technology use frustrating for the technologically perceptive, let alone the many teachers who may be somewhat techno-phobic (Ezziane, 2007).‖

1) *Individual Perceptions about ICTs*

One way to assess an individual's approach to computer use for instruction is by testing an individual's attitudes to this (Graff et al., 2001). Understanding learner perceptions of technology and its impact on their practice will help in addressing technology-training of the user (Zhao and Bryant, 2006). Learner attitudes are reportedly strongly related to their success in using technology (Bataineh& Abdel-Rahman, 2006). Students‘ use of computer and Internet depends on their perceived usefulness in terms of effective communication and access to information to complete projects and assignments efficiently (Gay et al., 2006). However, limited research has been published about students' perceptions of their computer literacy, particularly, in developing states (Bataineh& Abdel-Rahman, 2006).Technology paradigm shifts changed not only the way of computing but also how the technology itself is perceived by society (Ezziane, 2007)Educational technologies are generally perceived as a welcome addition to the pedagogical and learning tools (Sasseville, 2004). However, by compelling instructors to collaborate with people outside the classroom (government agencies, university administrators, technical support staff etc), technology can be perceived as a threat to the private practice of pedagogy (Aaron et al., 2004). The relevant concern, then, is how well teachers perceive and address the challenges for education (Knight et al., 2006). Based on the perceptual differences, Mehra&Mital (2007) have categorized learners into:

- Cynics: Those with negative perceptions about eLearning but strong pedagogical beliefs therefore unwilling to change beyond instrumental use of ICTs;
- Moderates: They like ICTs and are ready to change and adapt to new pedagogical practices with some guidance and training;
- Adaptors: These are the intellectual leaders who use eLearning for inner progress and external enhancements by continuously enriching their teaching and learning with leading-edge technologies.

2) *Organizational Perceptions/Approaches*

Aviram& Tami (2004) have extracted seven approaches: administrative, curricular, didactic, organizational, systemic, cultural and ideological and five attitudes: agnostic, conservative, moderate, radical, and extreme radical attitude towards the application of ICTs in HEIs (see Table 1 for details on these approaches. Administrative, Curricular, Didactic and Organizational approaches are more ―instrumental‘ than Systemic, Cultural and Ideological approaches, which emphasize broader and substantive view/role of ICTs in higher education. The instrumental view is mostly supported by the administrators, bureaucrats and politicians (Baumeister, 2006). While substantive approaches are possessed mostly by the academics and intellectuals who maintain that eLearning technologies must systematically change the educational culture according to the ideological requirements of a particular context (Mehra&Mital, 2007).

Table 1 Perceptions about the Organizational Roles of ICTs

1	Administrative	The availability of technology is the progress and an important aim, so focus is on the quantity and quality of equipment.
2	Curricular	The use of ICTs with a specific curricular aim. Technology is conceived as a neutral tool in the service of prevailing subject matters.
3	Didactic	Didactic approach dictates the inevitable or desirable change that can be brought through ICT in pedagogy.
4	Organizational	ICTs can help creating viable, flexible and robust organizational structures to teach, learn and administer effectively.
5	Systemic	ICTs have to be used systematically. All the changes must be preplanned and predefined.
6	Cultural	Cultural approach recognizes that the ICT revolution has powerful defining impact our culture and thus lives.
7	Ideological	Philosophical or critical social thinkers believe that whatever the change, it should be in tune with the Social-values of the society.

Adapted from: Aviram& Tami (2004)

Administrative, Curricular, Didactic and Organizational approaches are more _instrumental' than Systemic, Cultural and Ideological approaches, which emphasize broader _substantive view' or role of ICTs in higher education. The instrumental view is mostly supported by the administrators, bureaucrats and politicians (Baumeister, 2006). While substantive approaches are possessed mostly by the academics and intellectuals who maintain that eLearning technologies must systematically change the educational culture according to the ideological requirements of a particular context.

3) *Demographic Diversities*

Due to the demographic disparities, users hold different conceptions of ICTs and eLearning therefore express varying attitudes in the development and use of these tools. Given that the perceptions of every developer and user of ICTs vary (Sasseville, 2004), there is a multiplicity of user-theories forming a continuum of approaches about the nature and role of ICTs and attitudes about the extent of change required (Kopyc, 2007). Teachers, students and any other users of ICTs, behave according to their demographic characteristics of age, educational level, cultural background, physical and learning disabilities, experience, personal goals and attitudes, preferences, learning styles, motivation, reading and writing skills, computer skills, ability to work with diverse cultures, familiarity with differing instructional methods and previous experience with e-learning (Moolman&Blignaut, 2008).For example, male students prefer using computers in their learning than females. Individual differences are evident in terms of attitudes to computer-based learning and Internet use and that these differences exist principally on two levels, which are nationality and cognitive learning style (Graff et al., 2001). "Net Generation" is a force for educational transformation. They process information differently than previous generations, learn best in highly customizable environments, and look to teachers to create and structure their learning experience (Dinevski&Kokol, 2005) furthermore, male students have more positive perceptions about computers and information technology than female students. Older students may have a somewhat more positive perception of computers (Gay et al., 2006). Students bring prior knowledge to their learning experiences. This prior knowledge is known to affect how students encode and later retrieve new information learned (DiCerbo, 2007).

4) *Resistance to Change*

The user-resistance and reluctance to change is widely investigated topic in eLearning (see for example, Jager&Lokman, 1999; Sasseville, 2004; Loing, 2005; Vrana, 2007; Kanuka, 2007; Mehra&Mital, 2007). Since, teachers decide about what happens in the classroom therefore their acceptance plays a dominant role in the successful use of computers in the classroom (Aaron et al., 2004). Although most of the teachers have adopted ICTs like power point slides and internet into their teaching, they are still unwilling to adopt more sophisticated computer-based teaching innovations (Mehra&Mital, 2007)." It has been found that new things are intimidating and cause resistance (Jager&Lokman, 1999). For example, if teachers refuse to use ICTs in their classrooms, then eLearning can never progress except limited benefits. Furthermore, due to the innovative nature of ICT-enabled projects, the developers must have a keen understanding of the innovation process, identify the corresponding requirements for successful adoption, and harmonize plans and actions accordingly (Tinio, 2002). In Canada, teachers are reluctant to integrate technological innovations into their daily scholarly activities and, at least in Quebec, this situation has not really changed over the past few years (Sasseville, 2004).Within universities the decision makers and academics are sometimes reluctant to change curricula and pedagogic approaches; teaching staff and instructors lack incentive and rewards in a system where professional status and career trajectories are based on research results rather than pedagogic innovation (Loing, 2005). There are many obstacles for implementation of the ICT in universities. Some of them are classical, as are e.g. inertia of behavior of people, their resistance to changes, etc. If the ICT should

serve properly, it should enforce an order in all folds of the university life. People who loose their advantage of the better access to information have a fear from order. Regrettably, managers sometimes belong to this category (Vrana, 2007).Technological change is not perceived as a collective experience rather a personal challenge therefore, solutions to the problem of integrating technological innovations into the pedagogy are more focused on the individual teachers (Sasseville, 2004). Some teachers are strongly advocate the technological innovation but may resist in accepting technology as an integral part of the learning process. These divergent reactions and concerns have thus created a continuum that represents various attitudes towards technology (Juniu, 2005). Similarly, —nexperience may lead to developing learners' anxiety (Moolman&Blignaut, 2008)."

5) *Training Ineffectiveness*

The gap between user and ICTs is possible if user-training is not undertaken effectively. Almost every research recording the perceptions and attitudes of eLearning-users reports the dissatisfaction from the training facilities, contents and duration with regard to eLearning tools for teaching, learning and administrative purposes (see for example, Gray et al., 2003; Loing, 2005; Johnson et al., 2006; Wells, 2007; Mehra&Mital, 2007). Albion (1999) noted this some 18 years ago that —ascommunity expectations for integration of information technology into the daily practices of teaching grow, it will become increasingly important that all teachers are adequately prepared for this dimension of their professional practice."User training includes the training of both the developers or ICT-professionals and Non-ICT users. Both the groups need computer literacy of the levels of their requirements. —Alarge body of literature supports the idea that technology training is the major factor that could help teachers develop positive attitudes toward technology and integrating technology into curriculum (Zhao & Bryant, 2006). Teachers need training for technology-integration —ni curriculum areas that can be replicated in their own classrooms not training that focuses on software applications and skill development (Schou, 2006)." The developers need such _omputing-curriculum' which covers not only the technological aspects of computer hardware and software but also the human and organizational dimensions of these tools when placed in use.

V. Conclusions

Digital literacy is a universal issue for HEIs and researchers. The new ICTs are forcing academicians to postulate refined theories for learning (Oliver, 2002). Our culture is no longer literary and artistic only, —ti is also technologic and scientific (Sasseville, 2004)." The paradigm shift in HEIs refers not only to the departure from the traditional pedagogy, learning and education-management; it also features changes within eLearning environments for teaching, learning and administrative purposes (Young, 2003; Baumeister, 2006). This paradigm shift is described in terms of the progress in digital literacy from old-ICTs to new-ICTs in three stages of traditional-eLearning, blended-eLearning and contemporary virtual-eLearning.Furthermore, digital literacy of students is squarely mounted on the computer competencies of the teachers and academicians because students cannot acquire computer literacy without a —computer literate faculty (Johnson et al., 2006)." Thus, computer literacy is one of the most important skills in today's competitive environment therefore government and HEIs are required to provide technical and political support to the faculty for successfully passing on digital knowledge and skills (Ezziane, 2007).

VI. References

1) Aaron, M., Dicks, D., Ives, C. & Montgomery, B. (2004) —Planning for Integrating Teaching Technologies", Canadian Journal of Learning and Technology, 30(2), spring. [Online]. Available from: http://www.cjlt.ca/ (accessed 14 May, 2007).
2) Abrami, P. C., Bernard , R. M., Wade, A., Schmid, R. F., Borokhovski, E., Tamim, R., Surkes, M. A., Lowerison, G., Zhang, D., Nicolaidou, I., Newman, S., Wozney, l., and Peretiatkowicz, A. (2006), —A Review of e-Learning in Canada: A Rough Sketch of the Evidence, Gaps and Promising Directions", Canadian Journal of Learning and Technology, 32(3), Fall/Autumn. [Online]. Available from: http://www.cjlt.ca/ (accessed 14 May, 2007).
3) Albion, P. R (1999), —Sel-Efficacy Beliefs as an Indicator of Teachers' Preparedness for Teaching with Technology", [Online]. Available from: http://www.usq.edu.au/users/albion/papers /site99/1345.html (accessed 10 May, 2007).
4) Andriole, S. J. (2006), —Business Technology Education in the Early 21st Century: The Ongoing Quest for Relevance", Journal of Information Technology Education, 5. [Online]. Available from: http://jite.org/documents/Vol5/ (accessed 3 July, 2007).
5) Aviram, A. &Eshet-Alkalai, Y. (2006), —Towards a Theory of Digital Literacy: Three Scenarios for the Next Steps", European Journal of Open, Distance and E-Learning. [Online]. Available from: http://www.eurodl.org/ (accessed 11 May, 2007).
6) Aviram, R. & Tami, D. (2004), —The impact of ICT on education: the three opposed paradigms, the lacking discourse", [Online]. Available from: http://www.informatik.uni-bremen.de/~mueller/kr-004/ressources/ict_impact.pdf (accessed 13 July, 2007).
7) Barnes, K., Marateo, R. C., & Ferris, S. P. (2007), —Teaching and Learning with the Net Generation", Innovate Journal of Online Education, 3(4). Available from: http://Innovateonline.info (accessed 10 April, 2007).
8) Bataineh, R. F. &Bani-Abdel-Rahman, A. A. (2006), —Jordanian EFL students' perceptions of their computer literacy: An exploratory case study", International Journal of Education and Development using ICT, 2(2). Available from:

http://ijedict.dec.uwi.edu//viewarticle.php?id=169 &layout=html) accessed 10 April, 2007).

9) Baumeister, H. (2006), "Networked Learning in the Knowledge Economy - A Systemic Challenge for Universities", European Journal of Open, Distance and E-Learning. [Online]. Available from: http://www.eurodl.org/ (accessed 10 April, 2007).
10) Beebe, M. A. (2004), "Impact of ICT Revolution on the African Academic Landscape", In Proceedings of CODESRIA Conference on Electronic Publishing and Dissemination. Dakar, Senegal. 1 -2 September 2004. [Online]. Available from: http://www.codesria.org/Links/conferences/el_publ /beebe.pdf (accessed 5 April, 2007).
11) Bondarouk, T. V. (2006), "Action-oriented group learning in the implementation of information technologies: results from three case studies", European Journal of Information Systems, 15, pp. 42–53. [Online] Available from: http://www.palgrave-journals.com/ejis/ (accessed 10 April, 2007).
12) Buzhardt, J. &Heitzman-Powell, L. (2005), "Stop blaming the teachers: The role of usability testing in bridging the gap between educators and technology", Electronic Journal for the Integration of Technology in Education, 4(13). [Online] Available from: http://ejite.isu.edu/Volume3No1/ (accessed 10 April, 2007).
13) Chan, A. & Lee, M. J. W. (2007), "We Want to be Teachers, Not Programmers: In Pursuit of Relevance and Authenticity for Initial Teacher Education Students Studying an Information Technology Subject at an Australian University", Electronic Journal for the Integration of Technology in Education, 6, pp. 79. [Online] Available from: http://ejite.isu.edu/Volume3No1/ (accessed 9 April, 2007).
14) DiCerbo, K. E. (2007), "Knowledge Structures of Entering Computer Networking Students and Their Instructors", Journal of Information Technology Education, 6. [Online] Available from: http://jite.org/documents/Vol6/ (accessed 13 July, 2007).
15) Dinevski, D. &Kokol, D. P. (2005), "ICT and Lifelong Learning", European Journal of Open, Distance and E-Learning, [Online] Available from: http://www.eurodl.org/ (accessed 12 April, 2007).
16) Ekstrom, J. J., Gorka, S., Kamali, R., Lawson, E., Lunt, B., Miller, J. &Reichgelt, H. (2006), "The Information Technology Model Curriculum", Journal of Information Technology Education, 5. [Online] Available from: http://jite.org/documents/Vol5/ (accessed 13 July, 2007).
17) Ezer, J. (2006), "India and the USA: A Comparison through the Lens of Model IT Curricula", Journal of Information Technology Education, 5. [Online] Available from: http://jite.org/documents/Vol5/ (accessed 13 July, 2007).
18) Ezziane, Z. (2007), "Information Technology Literacy: Implications on Teaching and Learning", Journal of Educational Technology & Society, 10 (3), pp. 175-191. [Online] Available from: http://www.ask4research.info/ (accessed April 10, 2007).
19) Gay, G., Mahon, S., Devonish, D., Alleyne, P. &Alleyne, P. G. (2006), "Perceptions of information and communication technology among undergraduate management students in Barbados", International Journal of Education and Development using ICT, 2(4). [Online] Available from: http://ijedict.dec.uwi.edu// accessed 11 May, 2007.
20) Glogoff, S. (2005), "Instructional Blogging: Promoting Interactivity, Student-Centered Learning, and Peer Input", Innovate Journal of Online Education, 1(5), June/July. [Online] Available from: http://Innovateonline.info (accessed 10 April, 2007).
21) Graff, M., Davies, J. &McNorton, M. (2001), "Cognitive Style and Cross Cultural Differences in Internet Use and Computer Attitudes", European Journal of Open, Distance and E-Learning, [Online] Available from: http://www.eurodl.org/ (accessed 10 April, 2007).
22) Gray, D. E., Ryan, M. &Coulon, A. (2003), "The Training of Teachers and Trainers: Innovative Practices, Skills and Competencies in the use of eLearning", European Journal of Open, Distance and E-Learning. [Online] Available from: http://www.eurodl.org/ (accessed 9 April, 2007).
23) Hameed, T. (2007), "ICTas an enabler of socio-economic development", [Online] Available from: http://www.itu.int/osg/spu/digitalbridges/materials/ hameed-paper.pdf (accessed 24 June, 2007).
24) Higher Education Commission (HEC) (2008), "eReforms: PERN, PRR, eLearning, CMS & Digital Library", [Online] Available from: http://www.hec.gov.pk/new/eReforms/ eReforms.htm (accessed 14 July, 2007).
25) Jager, A. K. &Lokman, A. H. (1999), "Impacts of ICT in education. The role of the teacher and teacher training", In Proceedings of The European Conference on Educational Research, Lahti, Finland 22 - 25 September. Stoas Research, Wageningen, The Netherlands. Retrieved April 10, 2007, from
26) Johnson, D. W., Bartholomew, K. W. & Miller, D. (2006), "Improving Computer Literacy of Business Management Majors: A Case Study", Journal of Information Technology Education, 5. [Online] Available from: http://jite.org/documents/Vol5/ (accessed 14 July, 2007).
27) Juniu, S. (2005), "Digital Democracy in Higher Education Bridging the Digital Divide", Innovate Journal of Online Education, 2(1),

October/November. [Online] Available from: http://Innovateonline.info (accessed 10 April, 2007).
28) Kanuka, H. (2007), "Instructional Design and eLearning: A Discussion of Pedagogical Content Knowledge as a Missing Construct", e-Journal of Instructional Science and Technology, 9(2). [Online] Available from: http://www.usq.edu.au/electpub/e-jist/docs/vol9_no2/default.htm (accessed 18 July, 2007).
29) Klamma, R., Chatti, M. A., Duval, E., Hummel, H., Hvannberg, E. H., Kravcik, M., Law, E., Naeve, A., & Scott, P. (2007), "Social Software for Life-long Learning", Journal of Educational Technology & Society, 10(3), pp. 72-83. [Online] Available from: http://www.ask4research.info/ (accessed June 24, 2007).
30) Knight, C., Knight, B. A. &Teghe, D. (2006), "Releasing the pedagogical power of information and communication technology for learners: A case study", International Journal of Education and Development using ICT, 2(2). [Online] Available from: http://ijedict.dec.uwi.edu// (accessed 11 May, 2007).
31) Kopyc, S. (2007), "Enhancing Teaching with Technology: Are We There Yet?", Innovate Journal of Online Education, 3(2), December 2006/January 2007. [Online] Available from: http://Innovateonline.info (accessed 10 April, 2007).
32) LaCour, S. (2005), "The future of integration, personalization, and ePortfolio technologies", Innovate Journal of Online Education, 1(4), April/May. [Online] Available from: http://Innovateonline.info (accessed 10 April, 2007).
33) Loing, B. (2005), "ICTand Higher Education. General delegate of ICDE at UNESCO", In Processing of 9th UNESCO/NGO Collective Consultation on Higher Education, 6-8 April 2005. [Online] Available from: http://ong-comite-liaison.unesco.org/ongpho/acti/3/11/rendu/20/pdfen.pdf (accessed 24 June, 2007).
34) Luck, P. & Norton, B. (2005), "Problem Based Management Learning-Better Online?", European Journal of Open, Distance and E-Learning. [Online] Available from: http://www.eurodl.org/ (accessed 10 April, 2007).
35) Macleod, H. (2005), "What role can educational multimedia play in narrowing the digital divide?", International Journal of Education and Development using ICT, 1(4). [Online] Available from: http://ijedict.dec.uwi.edu// (accessed 11 May, 2007).
36) Manochehr, N. (2007), "The Influence of Learning Styles on Learners in E-Learning Environments: An Empirical Study", Computers in Higher Education and Economics Review, 18. [Online] Available from:
37) http://www.economicsnetwork.ac.uk/cheer.htm (accessed 11 April, 2007).
38) Martin, F. &Dunsworth, Q. (2007), "A Methodical Formative Evaluation of Computer Literacy Course: What and How to Teach", Journal of Information Technology Education, 6. [Online] Available from: http://jite.org/documents/Vol6/ (accessed 10 October, 2007).
39) Mathur, S. K. (2006), "Indian Information Technology Industry: Past, Present and Future& A Tool for National Development", Journal of Theoretical and Applied Information Technology. [Online] Available from: http://www.jatit.org (accessed 6 October, 2007).
40) Mehra, P. &Mital, M. (2007), "Integrating technology into the teaching-learning transaction: Pedagogical and technological perceptions of management faculty", International Journal of Education and Development using ICT, 3(1). [Online] Available from : http://ijedict.dec.uwi.edu// (accessed 10 October, 2007).
41) Mejias, U. (2006), "Teaching Social Software with Social Software", Innovate Journal of Online Education, 2(5), June/July. [Online] Available from: http://Innovateonline.info (accessed 9 April, 2007).
42) Moolman, H. B., &Blignaut, S. (2008), "Get set! e-Ready, … e-Learn! The e-Readiness of Warehouse Workers", Journal of Educational. Technology & Society, 11(1), pp. 168-182. [Online] Available from: http://www.ask4research.info/ (accessed 10 April, 2007).
43) Oliver, R. (2002), "The role of ICT in higher education for the 21st century: ICT as a change agent for education", [Online] Available from: http://elrond.scam.ecu.edu.au/oliver/2002/ he21.pdf (accessed 13 April, 2007).
44) Phillips, P., Wells, J., Ice, P., Curtis, R. & Kennedy, R. (2008), "A Case Study of the Relationship Between Socio-Epistemological Teaching Orientations and Instructor Perceptions of Pedagogy in Online Environments", Electronic Journal for the Integration of Technology in Education, 6, pp. 3-27. [Online] Available from: http://ejite.isu.edu/Volume6No1/ (accessed 7 April, 2007).
45) Radosevich, D. & Kahn, P. (2006), "Using Tablet Technology and Recording Software to Enhance Pedagogy", Innovate Journal of Online Education, 2(6), Aug/Sep. [Online] Available from: http://Innovateonline.info (accessed 8 April, 2007).
46) Russell, G. (2005), "TheDistancing Question in Online Education. Innovate Journal of Online Education, 1(4), April/May. [Online] Available from: http://Innovateonline.info (accessed 8 April, 2007).

47) Sahay, S. (2004), —Beyond utopian and nostalgic views of information technology and education: Implications for research and practice", Journal of the Association for Information Systems, 5(7), pp. 282-313. [Online] Available from: (accessed 7 April, 2007).
48) Sasseville, B. (2004), —Integrating Information and Communication Technology in the Classroom: A Comparative Discourse Analysis", Canadian Journal of Learning and Technology, 30(2), Spring. [Online] Available from: http://www.cjlt.ca/ (accessed 14 May, 2007).
49) Schou, S. B. (2006), —A Study of Student Attitudes and Performance in an Online Introductory Business Statistics Class", Electronic Journal for the Integration of Technology in Education, 6, pp. 71-78. [Online] Available from: http://ejite.isu.edu/Volume3No1/(accessed 9 April, 2007).
50) Sife, A. S., Lwoga, E. T. &Sanga, C. (2007), —New technologies for teaching and learning: Challenges for higher learning institutions in developing countries", International Journal of Education and Development using ICT, 3(1). [Online] Available from: http://ijedict.dec.uwi.edu// (accessed 21 July, 2007).
51) Sirkemaa, S. (2001), —Information technology in developing a meta-learning environment", European Journal of Open, Distance and E-Learning. [Online] Available from: http://www.eurodl.org/ (accessed 10 April, 2007).
52) Thompson, J. (2007), —Is Education 1.0 Ready for Web 2.0 Students?", Innovate Journal of Online Education, 3(4), April/May. [Online] Available from: http://Innovateonline.info (accessed 22 October, 2007).
53) Tinio, V. L. (2002), —ICT in education", In Proceedings of UNDP for the benefit of participants to the World Summit on the Information Society, UNDP's regional project, the Asia-Pacific Development Information Program (APDIP), in association with the secretariat of the Association of Southeast Asian Nations (ASEAN). [Online] Available from: http://www.apdip.net/publications/ iespprimers/eprimer-edu.pdf (accessed 14 July, 2007).
54) Vrana, I. (2007), —Changes required by ICT era are painful sometimes", In Proceedings of CAUSE98, an EDUCAUSE conference, [Online] Available from: http://www.educause.edu/copyright.html (accessed 10 October, 2007).
55) Ward, T., Monaghan. K. &Villing, R. (2006), —MyVLE: A case study in building a universal telematic education environment for a small university", European Journal of Open, Distance and E-Learning. [Online] Available from: http://www.eurodl.org/ (accessed 10 April, 2007).
56) Wells, R. (2007), —Challenges and opportunities in ICT educational development: A Ugandan case study", International Journal of Education and Development using ICT, 3(2). [Online] Available from: http://ijedict.dec.uwi.edu// (accessed 21 July, 2007).
57) Wijekumar, K. (2005), —Creating Effective Web-Based Learning Environments: Relevant Research and Practice", Innovate Journal of Online Education, 1(5), June/July. [Online] Available from: http://Innovateonline.info (accessed 10 April, 2007).
58) Wikipedia (2009), —eLearning", [Online] Available from: http://www.Wikipedia.org/ (accessed 10 February, 2009).
59) Wims, P. & Lawler, M. (2007), —Investing in ICTs in educational institutions in developing countries: An evaluation of their impact in Kenya", International Journal of Education and Development using ICT, 3(1). [Online] Available from: http://ijedict.dec.uwi.edu// (accessed 21 July, 2007).
60) Wood, R. E. (2004), —Scaling Up: From Web-Enhanced Courses to a Web-Enhanced Curriculum", Innovate Journal of Online Education, 1(1), Oct/Nov. [Online] Available from: http://Innovateonline.info (accessed 6 April, 2007).
61) Young, L. D. (2003), —Bridging Theory and Practice: Developing Guidelines to Facilitate the Design of Computer-based Learning Environments", Canadian Journal of Learning and Technology, 29(3), Fall/Autumn. [Online] Available from: http://www.cjlt.ca/ (accessed 13 May, 2007).
62) Zhao, Y. &LeAnna Bryant, F. (2006), —Can Teacher Technology Integration Training Alone Lead to High Levels of Technology Integration? A Qualitative Look at Teachers' Technology Integration after State Mandated Technology Training", Electronic Journal for the Integration of Technology in Education, 5, pp. 53-62. [Online] Available from: http://ejite.isu.edu/Volume5No1/ (accessed 9 April, 2007).

A Note on Intuitionistic Fuzzy Hypervector Spaces

Sanjay Roy, T. K. Samanta

GJCST Classification
I.2.3.I5.1

***Abstract*-The notion of Intuitionistic fuzzy hypervector space has been generalized and a few basic properties on this concept are studied. It has been shown that the intersection and union o fan arbitrary family of Intuitionistic fuzzy .**
Hypervector spaces are also Intuitionistic fuzzy hypervector space. Lastly, the no-tion of a linear transformation on a hypervector space is in- troduced and established an important theorem relative to In- tuitionistic fuzzy hypervector spaces.

Keywords-Intuitionistic fuzzy hypereld, Intuitionistic fuzzy hypervector spaces, Linear transformation.2010 Mathematics Subject Classi¯cation: 03F55, 08A72

I. Introduction

The notion of hyperstructure was introduced by F. Marty in 1934. Thenhe established the defnition of hypergroup [4] in 1935. Since then many re- searchers have studied and developed (for example see [5] , [6]) the concept of different types of hyperstructures in di®erent views. In 1990 M. S. Tallini[10] introduced the notion of hypervector spaces. Then in 2005 R. Ameri [1] also studied this spaces extensively. In our previous papers ([7], [8]), we also introduced the notion of a hypervector spaces in more general form than the previous concept of hypervector space and thereafter established a few useful theorems in this space.The concept of intuitionistic Fuzzy set, as a generalization of a fuzzy set was first introduced by Atanassov [3]. Then many researchers ([2], [9], [12]) applied this notion to norm, Continuity and Uniform Convergence etc. At the present time many researchers (for example [11]) are trying to apply this concept on the hyperstructure theory.In this paper, the concept of Intuitionistic fuzzy hypervector space is introduced and a few basic properties are developed. Further it has been shown that the intersection and union of a arbitrary family of Intuitionistic fuzzy hypervector spaces are also Intuitionistic fuzzy hypervector space. Lastly we have introduced the notion of a linear transformation on a hypervector space and established an important theorem relative to Intuitionistic fuzzy hypervectorspace.

II. Preliminaries

This section contain some basic definition and preliminaryresults whichwill be needed

Definition 2.1 *[6] A* ***hyperoperation*** *over a non empty set X is a mapping of $X\times X$ into the set of all non empty subsets of X.*

Definition 2.2 *[6] A non empty set R with exactly one hyperoperation $'\#'$ is a* ***hypergroupoid.***

Let $(X\ ,\ \#)$ be a hypergroupoid. For every point $x \in X$ and every non empty

A Note on Intuitionistic Fuzzy Hypervector Spaces 3

subset A of X , we defined $x\ \#\ A\ =\ \bigcup_{a\in A}\{x\ \#\ a\}$.

Definition 2.3 *[6] A hypergroupoid $(X\ ,\ \#)$ is called a* ***hypergroup*** *if*

(i) $x\ \#\ (y\ \#\ z) = (x\ \#\ y)\ \#\ z$

(ii) $\exists 0\ \in X$ *such that for every* $a \in X$ *, there is unique element* $b \in X$ *for which* $0 \in a\ \#\ b$ *and* $0 \in b\ \#\ a$ *. Here* b *is denoted by* $-a$ *.*

About[1]-Department of Mathematics, South BantraRamkrishna Institution , West Bengal, India. e-mail: sanjaypuremath@gmail.com
About[2]-Department of Mathematics, Uluberia College, West Bengal, India.e-mail: mumpu¡tapas5@yahoo.co.in

(iii) *For all* $a\ ,\ b\ ,\ c \in X$ *if* $a \in b\ \#\ c\ ,$ *then* $b \in\ a\ \#\ (-c).$

Proposition 2.4 *[6]* (i) *In a hypergroup* $(X\ ,\ \#)\ ,\ -(-a) = a\ ,\ \forall a\ \in X.$

$(ii)\ 0\ \#\ a\ =\ \{a\}\ ,\ \forall a \in X\ ,$ *if* $(X\ ,\ \#)$ *is a commutative hypergroup.*

(iii) *In a commutative hypergroup* $(X\ ,\ \#)\ ,\ 0$ *is unique.*

Definition 2.5 *[6] A* ***hyperring*** *is a non empty set equipped with a hyper-*
ıddition '#' and a multiplication '.' such that $(X\ ,\ \#)$ *is a commutative hyper-*
ɹroup and $(X\ ,.)$ *is a semigroup and the multiplication is distributive across the*
ʊyperaddition both from the left and from the right and $a.0 = 0.a = 0,\ \forall\ a \in$
$X\ ,$ *where* 0 *is the zero element of the hyperring.*

Definition 2.6 *[6] A* ***hyperfield*** *is a non empty set* X *equipped with a hyperaddition '#' and a multiplication '.'such that*

(i) $(X\ ,\ \#\ ,\ .)$ *is a hyperring.*

(ii) $\exists$ *an element* $1 \in X,$ *called the identity element such that* $a.1 = a,$
$\forall\ a \in X$

(iii) *For each non zero element* a *in* $X\ ,\ \exists$ *an element* a^{-1} *such that* $a.a^{-1}=1$

(iv) $a.b\ =\ b.a\ ,\ \forall\ a\ ,\ b \in X.$

Definition 2.7 *[8] Let* $(F\ ,\oplus\ ,.)$ *be a hyperfield and* $(V\ ,\ \#)$ *be an additive commutative hypergroup . Then* V *is said to be a* ***hypervector space*** *over the hyperfield* F *if there exist a hyperoperation* $*:\ F\ \times\ V\ \rightarrow\ P^{*}\ (V\ )$ *such that*

(i) $a\ *\ (\alpha\ \#\ \beta)\ \subseteq\ a\ *\ \alpha\ \#\ a\ *\ \beta\ ,\qquad \forall\ a \in F\ and\ \forall\ \alpha\ ,\beta \in V$

(ii) $(a\ \oplus\ b)\ *\ \alpha \subseteq a\ *\ \alpha\ \#\ b\ *\ \alpha\ ,\qquad \forall\ a, b \in F\ and\ \forall\ \alpha \in V$

(iii) $(a\ .\ b)\ *\ \alpha\ =\ a\ *\ (b\ *\ \alpha)\ ,\qquad \forall\ a, b \in F\ and\ \forall\ \alpha \in V.$

(iv) $(-a)\ *\ \alpha\ =\ a\ *\ (-\alpha),\qquad \forall\ \alpha \in\ V\ and\ \forall\ a \in F.$

(v) $\alpha \in 1_F\ *\ \alpha,\ \theta \in\ 0\ *\ \alpha\ and\ 0\ *\ \theta\ =\ \theta,\qquad \forall\ \alpha \in V.$

Definition 2.8 *[1] Let* $f\ :\ X\ \rightarrow\ Y$ *be a mapping and* $\nu \in FS(Y).$ *Then we define* $f^{-1}(\nu) \in FS(X)$ *as follows:*

$f^{-1}(\nu)(x)\ =\ \nu(f(x)),\ \forall\ x \in X.$

Definition 2.9 *[12] Let E be a any set. An Intuitionistic fuzzy set (IFS) A of E is an object of the form $A = \{ (x, \mu_A(x), \nu_A(x)) : x \in E \}$, where the functions $\mu_A : E \to [0, 1]$ and $\nu_A : E \to [0, 1]$ denotes the degree of membership and the non-membership of the element $x \in E$ respectively and for every $x \in E$, $0 \leq \mu_A(x) + \nu_A(x) \leq 1$.*

III. INTUITIONISTIC FUZZY HYPERVECTOR SPACE

In this section we established the definition of intuitionisticfuzzyhypervector Spaces and deduce some important theorems.

Definition 3.1 *Let $(F, \oplus, .)$ be a hyperfield. An* ***intuitionistic fuzzy hyperfield*** *on F is an object of the form $A = \{ (a, \mu_F(a), \nu_F(a)) : a \in F \}$ satisfies the following conditions :*

(i) $\bigwedge_{x \in a \oplus b} \mu_F(x) \geq \mu_F(a) \wedge \mu_F(b)$, $\forall\ a,\ b \in F$

(ii) $\mu_F(-a) \geq \mu_F(a)$, $\forall\ a \in F$

(iii) $\mu_F(a.b) \geq \mu_F(a) \wedge \mu_F(b)$ $\forall\ a,\ b \in F$

(iv) $\mu_F(a^{-1}) \geq \mu_F(a)$, $\forall\ a (\neq 0) \in F$

(v) $\bigvee_{x \in a \oplus b} \nu_F(x) \leq \nu_F(a) \vee \nu_F(b)$, $\forall\ a,\ b \in F$

(vi) $\nu_F(-a) \leq \nu_F(a)$, $\forall\ a \in F$

(vii) $\nu_F(a.b) \leq \nu_F(a) \vee \nu_F(b)$, $\forall\ a,\ b \in F$

$(viii)$ $\nu_F(a^{-1}) \leq \nu_F(a)$, $\forall\ a (\neq 0) \in F$.

Result 3.2 If A is a intuitionistic fuzzy hyperfield of F, then

(i) $\mu_F(0) \geq \mu_F(a)$, $\forall\ a \in F$

(ii) $\mu_F(1) \geq \mu_F(a)$, $\forall\ a \in F \setminus \{0\}$

(iii) $\mu_F(0) \geq \mu_F(1)$

(iv) $\nu_F(0) \leq \nu_F(a)$, $\forall\ a \in F$

(v) $\nu_F(1) \leq \nu_F(a)$, $\forall\ a \in F \setminus \{0\}$

(vi) $\nu_F(0) \leq \nu_F(1)$

Proof : Obvious.

Definition 3.3 *Let $(V, \#, *)$ be a hypervector space over a hyperfield $(F, \oplus, .)$ and A be a intuitionistic fuzzy hyperfield in F. A intuitionistic fuzzy subset $B = \{ (x, \mu_V(x), \nu_V(x)) : x \in V \}$ of V is said to be a* ***intuitionistic fuzzy hypervector space*** *of V over a intuitionistic fuzzy hyperfield A, if the following conditions are satisfied:*

(i) $\bigwedge_{\alpha \in x\#y} \mu_V(\alpha) \geq \mu_V(x) \wedge \mu_V(y), \forall x, y \in V$

(ii) $\mu_V(-x) \geq \mu_V(x), \forall x \in V$

(iii) $\bigwedge_{y \in a*x} \mu_V(y) \geq \mu_V(x) \wedge \mu_F(a), \ \forall a \in F$ *and* $\forall x \in V$

(iv) $\mu_F(1) \geq \mu_V(\theta)$, *where θ be the null vector of V.*

(v) $\bigvee_{\alpha \in x\#y} \nu_V(\alpha) \leq \nu_V(x) \vee \nu_V(y), \forall x, y \in V$

(vi) $\nu_V(-x) \leq \nu_V(x), \forall x \in V$

(vii) $\bigvee_{y \in a*x} \nu_V(y) \leq \nu_V(x) \vee \nu_F(a), \forall a \in F$ *and* $\forall x \in V$

$(viii)$ $\nu_F(1) \leq \nu_V(\theta)$, *where θ be the null vector of V.*

Here we say that B is a intuitionistic fuzzy hypervector space over a intuitionistic fuzzy hyperfield A.
Result 3.4 If B is a intuitionistic fuzzy hypervector space over a intuition-istic fuzzy hyperfield A, then

(i) $\mu_F(0) \geq \mu_V(\theta)$,

(ii) $\mu_V(\theta) \geq \mu_V(x), \forall x \in V$

(iii) $\mu_F(0) \geq \mu_V(x), \forall x \in V$

(iv) $\nu_F(0) \leq \nu_V(\theta)$,

(v) $\nu_V(\theta) \leq \nu_V(x), \forall x \in V$

(vi) $\nu_F(0) \leq \nu_V(x), \forall x \in V$

Proof : Obvious.

Theorem 3.5 *Let V be a hypervector space over a hyperfield F and A be intuitionistic fuzzy hyperfield. Let $B \in IFS(V)$. Then B is a intuitionistic zzy hypervector space over A iff*

) $\bigwedge_{z \in a*x\#b*y} \mu_V(z) \geq (\mu_F(a) \wedge \mu_V(x)) \wedge (\mu_F(b) \wedge \mu_V(y))$, $\forall x, y \in V$ *and* $a, b \in F$

i) $\mu_F(1) \geq \mu_V(\theta)$ *where θ be the null vector of V.*

ii) $\bigvee_{z \in a*x\#b*y} \nu_V(z) \leq (\nu_F(a) \vee \nu_V(x)) \vee (\nu_F(b) \vee \nu_V(y))$, $\forall x, y \in V$ *and* $a, b \in F$

v) $\nu_F(1) \leq \nu_V(\theta)$ *where θ be the null vector of V.*

Proof: First we suppose that B is a intuitionistic fuzzy hypervector space over the intuitionistic fuzzy hyperfield A. Then for a, b $\in$ F and $x, y \in$ V, we have

$$\begin{aligned}\bigwedge_{z\in a*x\#b*y}\mu_V(z) &= \bigwedge_{z\in\alpha\#\beta,\alpha\in a*x,\beta\in b*y}\mu_V(z)\\ &=\bigwedge_{\alpha\in a*x,\beta\in b*y}(\bigwedge_{z\in\alpha\#\beta}\mu_V(z))\\ &\geq \bigwedge_{\alpha\in a*x,\beta\in b*y}(\mu_V(\alpha)\wedge\mu_V(\beta))\\ &= (\bigwedge_{\alpha\in a*x}\mu_V(\alpha))\bigwedge(\bigwedge_{\beta\in b*y}\mu_V(\beta))\\ &\geq (\mu_F(a)\wedge\mu_V(x))\bigwedge(\mu_F(b)\wedge\mu_V(y))\end{aligned}$$

A Note on Intuitionistic Fuzzy Hypervector Spaces

$$\begin{aligned}\bigvee_{z\in a*x\#b*y}\nu_V(z) &= \bigvee_{z\in\alpha\#\beta,\alpha\in a*x,\beta\in b*y}\nu_V(z)\\ &=\bigvee_{\alpha\in a*x,\beta\in b*y}(\bigvee_{z\in\alpha\#\beta}\nu_V(z))\\ &\leq \bigvee_{\alpha\in a*x,\beta\in b*y}(\nu_V(\alpha)\vee\nu_V(\beta))\\ &= (\bigvee_{\alpha\in a*x}\nu_V(\alpha))\bigvee(\bigvee_{\beta\in b*y}\nu_V(\beta))\\ &\leq (\nu_F(a)\vee\nu_V(x))\bigvee(\nu_F(b)\vee\nu_V(y))\end{aligned}$$

The second and fourth inequalities are directly follow.

conversely suppose that the inequalities of the theorem hold for all $x, y \in$ V and $\forall$ a, b $\in$ F.

$$\begin{aligned}\text{Then } \bigwedge_{z\in 1*x\#1*y}\mu_V(z) &\geq (\mu_F(1)\wedge\mu_V(x))\bigwedge(\mu_F(1)\wedge\mu_V(y))\\ &\geq (\mu_V(\theta)\wedge\mu_V(x))\bigwedge(\mu_V(\theta)\wedge\mu_V(y))\\ &= \mu_V(x)\wedge\mu_V(y)\end{aligned}$$

i.e $\bigwedge_{z\in x\#y}\mu_V(z) \geq \mu_V(x)\wedge\mu_V(y)$, as $x \in 1*x$ and $y \in 1*y$

$$\begin{aligned}\mu_V(-x) &= \bigwedge_{z\in -1*x}\mu_V(z)\text{ , as } -x\in -1*x\\ &\geq \bigwedge_{z\in -1*x\#0*x}\mu_V(z)\geq(\mu_F(-1)\wedge\mu_V(x))\bigwedge(\mu_F(0)\wedge\mu_V(x))\\ &\geq (\mu_F(1)\wedge\mu_V(x))\bigwedge\mu_V(x)\\ &\geq (\mu_V(\theta)\wedge\mu_V(x))\bigwedge\mu_V(x)\\ &= \mu_V(x)\wedge\mu_V(x)\\ &= \mu_V(x)\end{aligned}$$

i.e $\mu_V(-x) \geq \mu_(x)$

$$\bigwedge_{y\in a*x} \mu_V(y) \geq \bigwedge_{y\in a*x\#0*x} \mu_V(y) \geq (\mu_F(a) \wedge \mu_V(x)) \bigwedge (\mu_F(0) \wedge \mu_V(x))$$
$$= (\mu_V(x) \wedge \mu_F(a)) \bigwedge \mu_V(x) \text{ , as } \mu_F(0) \geq \mu_V(x)$$
$$= \mu_V(x) \wedge \mu_F(a)$$

The fourth inequality of definition 3.3 is directly follows.

Next $\bigvee_{z\in 1*x\#1*y} \nu_V(z) \leq (\nu_F(1) \vee \nu_V(x)) \bigvee (\nu_F(1) \vee \nu_V(y))$
$$= (\nu_V(\theta) \vee \nu_V(x)) \bigvee (\nu_V(\theta) \vee \nu_V(y))$$
$$= \nu_V(x) \vee \nu_V(y)$$

i.e $\bigvee_{z\in x\#y} \nu_V(z) \leq \nu_V(x) \vee \nu_V(y)$, as $x \in 1 * x$ and $y \in 1 * y$

$$\nu_V(-x) \leq \bigvee_{z\in -1*x} \nu_V(z) \text{ , as } -x \in -1 * x$$
$$\leq \bigvee_{z\in -1*x\,\#\,0*x} \nu_V(z)$$
$$\leq (\nu_F(-1) \vee \nu_V(x)) \vee (\nu_F(0) \vee \nu_V(x))$$
$$\leq (\nu_F(1) \vee \nu_V(x)) \bigvee \nu_V(x)$$
$$\leq (\nu_V(\theta) \vee \nu_V(x)) \bigvee \nu_V(x)$$
$$= \nu_V(x) \vee \nu_V(x)$$
$$= \nu_V(x)$$

i.e $\nu_V(-x) \leq \nu_(x)$

$$\bigvee_{y\in a*x} \nu_V(y) \leq \bigvee_{y\in a*x\#\,0*x} \nu_V(y)$$
$$\leq (\nu_F(a) \vee \nu_V(x)) \vee (\nu_F(0) \vee \nu_V(x))$$
$$= (\nu_V(x) \vee \nu_F(a)) \bigvee \nu_V(x) \text{ , as } \nu_F(0) \leq \nu_V(x)$$
$$= \nu_V(x) \vee \nu_F(a)$$

The eighth inequality of definition 3.3 is obvious.Therefore B is a intuitionistic fuzzy hypervector space over A.This completes the proof.

Definition 3.6 *Let* $B^{\alpha}_{(\alpha\in\Lambda)} = \{ x, \mu^{\alpha}_V(x), \nu^{\alpha}_V(x) : x \in V \}$ *be a family of ntuitionistic fuzzy hypervector spaces of a hypervector space V over the same ntuitionistic fuzzy hyperfield* $A = \{ (x, \mu_F(x), \nu_F(x)) : x \in F \}$. *Then he* ***intersection*** *of those intuitionistic fuzzy hypervector spaces is defined as*

$\bigcap_{\alpha\in\Lambda} B^{\alpha})(x) = \{ x, \bigwedge_{\alpha\in\Lambda} \mu^{\alpha}_V(x), \bigwedge_{\alpha\in\Lambda} \nu^{\alpha}_V(x) : \forall x \in V \}$

nd the ***union*** *of those intuitionistic fuzzy hypervector spaces is defined as*

$\bigcup_{\alpha\in\Lambda} B^{\alpha})(x) = \{ x, \bigvee_{\alpha\in\Lambda} \mu^{\alpha}_V(x), \bigvee_{\alpha\in\Lambda} \nu^{\alpha}_V(x) : \forall x \in V \}$

Theorem 3.7 The intersection of any family of intuitionistic fuzzy hypervector spaces of a hypervector space V is a intuitionistic fuzzy hypervectorspace.

A Note on Intuitionistic Fuzzy Hypervector Spaces

Proof: Let $\{B^\alpha : \alpha \in \Lambda\}$ be a family of intuitionistic fuzzy hypervector spaces of V over the same intuitionistic fuzzy hyperfield $A = \{(x, \mu_F(x), \nu_F(x)) : x \in F\}$.

Let $B = (\bigcap_{\alpha\in\Lambda} B^\alpha)(x) = \{(x, \mu_V(x), \nu_V(x)) : x \in V\}$

where $\mu_V(x) = \bigwedge_{\alpha\in\Lambda} \mu_V^\alpha(x)$ and $\nu_V(x) = \bigwedge_{\alpha\in\Lambda} \nu_V^\alpha(x)$

Let $x, y \in V$ and a, b $\in$ F

$$\begin{aligned}\bigwedge_{z\in a*x\#b*y} \mu_V(z) &= \bigwedge_{z\in a*x\#b*y}(\bigwedge_{\alpha\in\Lambda} \mu_V^\alpha(z)) \\ &= \bigwedge_{\alpha\in\Lambda}(\bigwedge_{z\in a*x\#b*y} \mu_V^\alpha(z) \\ &\geq \bigwedge_{\alpha\in\Lambda}\{(\mu_F(a) \wedge \mu_V^\alpha(x)) \bigwedge(\mu_F(b) \wedge \mu_V^\alpha(y))\} \\ &= \{(\mu_F(a) \bigwedge(\bigwedge_{\alpha\in\Lambda} \mu_V^\alpha(x))\} \bigwedge\{\mu_F(b) \bigwedge(\bigwedge_{\alpha\in\Lambda} \mu_V^\alpha(y))\} \\ &= (\mu_F(a) \wedge \mu_V(x)) \bigwedge(\mu_F(b) \wedge \mu_V(y))\end{aligned}$$

Therefore $\bigwedge_{z\in a*x\#b*y} \mu_V(z) \geq (\mu_F(a) \wedge \mu_V(x)) \bigwedge(\mu_F(b) \wedge \mu_V(y))$

Again $\mu_F(1) \geq \mu_V^\alpha(\theta)$, $\forall\alpha \in \Lambda$

Therefore $\mu_F(1) \geq \bigwedge_{\alpha\in\Lambda} \mu_V^\alpha(\theta)$

i.e $\mu_F(1) \geq \mu_V(\theta)$

Next
$$\begin{aligned}\bigvee_{z\in a*x\#b*y} \nu_V(z) &= \bigvee_{z\in a*x\#b*y}(\bigwedge_{\alpha\in\Lambda} \nu_V^\alpha(z)) \\ &\leq \bigwedge_{\alpha\in\Lambda}(\bigvee_{z\in a*x\#b*y} \nu_V^\alpha(z)) \\ &\leq \bigwedge_{\alpha\in\Lambda}\{(\nu_F(a) \vee \nu_V^\alpha(x)) \bigvee(\nu_F(b) \vee \nu_V^\alpha(y))\} \\ &= \{(\nu_F(a) \bigvee(\bigwedge_{\alpha\in\Lambda} \nu_V^\alpha(x))\} \bigvee\{\nu_F(b) \bigvee(\bigwedge_{\alpha\in\Lambda} \nu_V^\alpha(y))\} \\ &= (\nu_F(a) \vee \nu_V(x)) \bigvee(\nu_F(b) \vee \nu_V(y))\end{aligned}$$

Therefore $\bigvee_{z\in a*x\#b*y} \nu_V(z) \leq (\nu_F(a) \vee \nu_V(x)) \bigvee(\nu_F(b) \vee \nu_V(y))$

Again $\nu_F(1) \leq \nu_V^\alpha(\theta)$, $\forall\alpha \in \Lambda$

Therefore $\nu_F(1) \leq \bigwedge_{\alpha\in\Lambda} \nu_V^\alpha(\theta)$

i.e $\nu_F(1) \leq \nu_V(\theta)$

Therefore B is also a intuitionistic fuzzy hypervector space over A.

This completes the proof.

Theorem 3.8 The union of any family of intuitionistic fuzzy hypervectorspaces of a hypervector space V is a intuitionistic fuzzy hypervector space.

Proof: Let $\{B^\alpha : \alpha \in \Lambda\}$ be a family of intuitionistic fuzzy hypervector spaces of V over the same intuitionistic fuzzy hyperfield $A = \{(x, \mu_F(x), \nu_F(x)) : x \in F\}$.

Let $B = \bigvee_{\alpha\in\Lambda} \mu_V^\alpha(x) = \{(x, \mu_V(x), \nu_V(x)) : x \in V\}$

where $\mu_V(x) = \bigvee_{\alpha\in\Lambda} \mu_V^\alpha(x)$ and $\nu_V(x) = \bigvee_{\alpha\in\Lambda} \nu_V^\alpha(x)$

Let $x, y \in$ V and a, b $\in$ F

$$\begin{aligned}
\bigwedge_{z\in a*x\#b*y} \mu_V(z) &= \bigwedge_{z\in a*x\#b*y}(\bigvee_{\alpha\in\Lambda} \mu_V^\alpha(z)) \\
&\geq \bigvee_{\alpha\in\Lambda}(\bigwedge_{z\in a*x\#b*y} \mu_V^\alpha(z)) \\
&\geq \bigvee_{\alpha\in\Lambda}\{(\mu_F(a) \wedge \mu_V^\alpha(x)) \bigwedge(\mu_F(b) \wedge \mu_V^\alpha(y))\} \\
&= \{(\mu_F(a) \bigwedge(\bigvee_{\alpha\in\Lambda} \mu_V^\alpha(x))\} \bigwedge\{\mu_F(b) \bigwedge(\bigvee_{\alpha\in\Lambda} \mu_V^\alpha(y))\} \\
&= (\mu_F(a) \wedge \mu_V(x)) \bigwedge(\mu_F(b) \wedge \mu_V(y))
\end{aligned}$$

Therefore $\bigwedge_{z\in a*x\#b*y} \mu_V(z) \geq (\mu_F(a) \wedge \mu_V(x)) \bigwedge(\mu_F(b) \wedge \mu_V(y))$

Again $\mu_F(1) \geq \mu_V^\alpha(\theta)$, $\forall\alpha \in \Lambda$

Therefore $\mu_F(1) \geq \bigvee_{\alpha\in\Lambda} \mu_V^\alpha(\theta)$

i.e $\mu_F(1) \geq \mu_V(\theta)$

$$\begin{aligned}
\textbf{Next } \bigvee_{z\in a*x\#b*y} \nu_V(z) &= \bigvee_{z\in a*x\#b*y}(\bigvee_{\alpha\in\Lambda} \nu_V^\alpha(z)) \\
&= \bigvee_{\alpha\in\Lambda}(\bigvee_{z\in a*x\#b*y} \nu_V^\alpha(z)) \\
&\leq \bigvee_{\alpha\in\Lambda}\{(\nu_F(a) \vee \nu_V^\alpha(x)) \bigvee(\nu_F(b) \vee \nu_V^\alpha(y))\} \\
&= \{(\nu_F(a) \bigvee(\bigvee_{\alpha\in\Lambda} \nu_V^\alpha(x))\} \bigvee\{\nu_F(b) \bigvee(\bigvee_{\alpha\in\Lambda} \nu_V^\alpha(y))\} \\
&= (\nu_F(a) \vee \nu_V(x)) \bigvee(\nu_F(b) \vee \nu_V(y))
\end{aligned}$$

Therefore $\bigvee_{z\in a*x\#b*y} \nu_V(z) \leq (\nu_F(a) \vee \nu_V(x)) \bigvee(\nu_F(b) \vee \nu_V(y))$

Again $\nu_F(1) \leq \nu_V^\alpha(\theta)$, $\forall\alpha \in \Lambda$

Therefore $\nu_F(1) \leq \bigvee_{\alpha\in\Lambda} \nu_V^\alpha(\theta)$

i.e $\nu_F(1) \leq \nu_V(\theta)$

Therefore B is also a intuitionistic fuzzy hypervector space over A.

A Note on Intuitionistic Fuzzy HypervectorSpacesThis completes the proof.

IV. LINEAR TRANSFORMATION

In this section we established the definition of a Linear transformation onhypervector Spaces and deduce a important theorem relative to a intuitionistic fuzzy concept.

Definition 4.1 *Let* $(V, \#, *)$ *and* $(W, \#', *')$ *be two hypervector space over the same hyperfield* $(F, \oplus, .)$. *A mapping* $T : V \to W$ *is called Linear transformation iff*

(*i*) $T(x \# y) \subseteq T(x) \#' T(y)$,

(*ii*) $T(a * x) \subseteq a *' T(x), \quad \forall x, y \in V$ *and* $a \in F$

(*iii*) $T(\theta) = \theta'$

Theorem 4.2 *Let* $(V, \#, *)$ *and* $(W, \#', *')$ *be two hypervector space over the same hyperfield* $(F, \oplus, .)$ *and* $T : V \to W$ *be a linear transformation. Let* $B = \{(y, \mu_W(y), \nu_W(y)) : y \in W\}$ *be a intuitionistic fuzzy hypervector space over* $A = \{(a, \mu_F(a), \nu_F(a)) : a \in F\}$. *Then* $T^{-1}(B) = \{x, T^{-1}(\mu_W)(x), T^{-1}(\nu_W)(x) : x \in V\}$ *is a intuitionistic fuzzy hypervector space of* V *over* A.

Proof: Let a, b ∈ F and $\alpha, \beta \in$ V.

Then $\bigwedge_{x \in a*\alpha \# b*\beta} T^{-1}(\mu_W)(x) = \bigwedge_{x \in a*\alpha \# b*\beta} \mu_W(T(x))$

$$= \bigwedge_{T(x) \in T(a*\alpha \# b*\beta)} \mu_W(T(x))$$

$$\geq \bigwedge_{T(x) \in a*'T(\alpha) \#' b*'T(\beta)} \mu_W(T(x))$$

$$\geq (\mu_F(a) \wedge \mu_W(T(\alpha)) \bigwedge (\mu_F(b) \wedge \mu_W(T(\beta)))$$

$$= (\mu_F(a) \wedge T^{-1}(\mu_W)(\alpha)) \bigwedge (\mu_F(b) \wedge T^{-1}(\mu_W)(\beta))$$

Again $T^{-1}(\mu_W)(\theta) = \mu_W(T(\theta)) = \mu_W(\theta')$, where θ' be a null vector of W.

$$\leq \mu_F(1)$$

i.e $\mu_F(1) \geq T^{-1}(\mu_W)(\theta)$

And $\bigvee_{x \in a*\alpha \# b*\beta} T^{-1}(\nu_W)(x) = \bigvee_{x \in a*\alpha \# b*\beta} \nu_W(T(x))$

$$= \bigvee_{T(x) \in T(a*\alpha \# b*\beta)} \nu_W(T(x))$$

$$\leq \bigvee_{T(x) \in a*'T(\alpha) \#' b*'T(\beta)} \nu_W(T(x))$$

$$\leq (\nu_F(a) \vee \nu_W(T(\alpha)) \bigvee (\nu_F(b) \vee \nu_W(T(\beta)))$$

$$= (\nu_F(a) \vee T^{-1}(\nu_W)(\alpha)) \bigvee (\nu_F(b) \vee T^{-1}(\nu_W)(\beta))$$

Again $T^{-1}(\nu_W)(\theta) = \nu_W(T(\theta)) = \nu_W(\theta')$, where θ' be a null vector of W.

$$\geq \nu_F(1)$$

i.e $\nu_F(1) \leq T^{-1}(\nu_W)(\theta)$

Therefore $T^{-1}(B)$ is a intuitionistic fuzzy hypervector space of a hypervector space V over A.

V. References

1) AMERI R. Fuzzy Hypervector Spaces Over Valued Field, Iranian J. of FuzzySystems Vol. 2, No. 1, (2005)pp. 37-47.
2) Dinda B, Samanta T. K. Intuitionistic Fuzzy Continuity and Uniform Convergence, Int. J. Open Problems Compt.Math.,Vol 3,No.1 (2010) 8 - 26.
3) K. AtanassovIntuitionistic fuzzy sets , Fuzzy Sets and Systems 20 (1986)87 - 96.
4) MARTY F. Role de la notion de hypergroupedans 1' elude de groupes nonabelians, comptesRenclus Acad. Sci. Paris, 201, 636-638,(1935).A Note on Intuitionistic Fuzzy Hypervector Spaces 13
5) MARTY F. Sur les groupes et les hypergroupes, attaches a une fractionrationnelle, Ann. Sci. Ecole Norm. Sup. (3) 53, 82-123, (1936).
6) Nakassis A. Expository and survey article recent results in hyperring andhyperfield theory, internet. J. Math. and Math. Sci.,11(2)(1988)209 - 220.
7) Roy S, Samanta T. K. A note on Hypervector Spaces ,(communicated)
8) Roy S, Samanta T. K. Hyper innerproduct Spaces ,(communicated)
9) Samanta T. K, Jebril IH. Finite dimentional intuitionistic fuzzy normedlinear space, Int. J. Open Problems Compt. Math., Vol 2, No. 4 (2009)574-591.
10) Tallini M S. Hypervector Spaces, Proceedings of fourth Int. Congress inAlgebraic Hyperstructures and Applications, Xanthi,Greece, World Scientifc(1990)167 - 174.
11) Torkzadeh, L. - Abbasi, M. - Zahedi, M.M. Some results of intuitionisticfuzzy weak dual hyper K-ideals, Iranian J. of Fuzzy Systems, Vol. 5, No. 1.,(2008), 65-78.
12) Vijayabalaji S, Thillaigovindan N, Jun YB. Intuitionistic Fuzzy n-normed linear space , Bull. Korean Math. Soc. 44 (2007) 291 - 308.

Global Journals Guidelines Handbook 2010

www.GlobalJournals.org

FELLOW OF INTERNATIONAL CONGRESS OF COMPUTER SCIENCE AND TECHNOLOGY (FICCT)

- FICCT' title will be awarded to the person after approval of Editor-in-Chief and Editorial Board. The title 'FICCT" can be added to name in the following manner e.g. **Dr. Andrew Knoll, Ph.D., FICCT, Er. Pettor Jone, M.E., FICCT**
- FICCT can submit two papers every year for publication without any charges. The paper will be sent to two peer reviewers. The paper will be published after the acceptance of peer reviewers and Editorial Board.
- Free unlimited Web-space will be allotted to 'FICCT 'along with subDomain to contribute and partake in our activities.
- A professional email address will be allotted free with unlimited email space.
- FICCT will be authorized to receive e-Journals - GJCST for the Lifetime.
- FICCT will be exempted from the registration fees of Seminar/Symposium/Conference/Workshop conducted internationally of GJCST (FREE of Charge).
- FICCT will be an Honorable Guest of any gathering hold.

ASSOCIATE OF INTERNATIONAL CONGRESS OF COMPUTER SCIENCE AND TECHNOLOGY (AICCT)

- AICCT title will be awarded to the person/institution after approval of Editor-in-Chef and Editorial Board. The title 'AICCTcan be added to name in the following manner:
  eg. **Dr. Thomas Herry, Ph.D., AICCT**
- AICCT can submit one paper every year for publication without any charges. The paper will be sent to two peer reviewers. The paper will be published after the acceptance of peer reviewers and Editorial Board.
- Free 2GB Web-space will be allotted to 'FICCT' along with subDomain to contribute and participate in our activities.
- A professional email address will be allotted with free 1GB email space.
- AICCT will be authorized to receive e-Journal GJCST for lifetime.
- A professional email address will be allotted with free 1GB email space.
- AICHSS will be authorized to receive e-Journal GJHSS for lifetime.

ANNUAL MEMBER

- Annual Member will be authorized to receive e-Journal GJCST 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) Register yourself using top right corner of Home page then Login from same place twice. If you are already registered, then login using your username and password.

(II) Choose corresponding Journal from "Research Journals" Menu.

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

(B) If you are using Internet Explorer (Although Mozilla Firefox is preferred), 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 as an attachment.

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

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 Times New Roman.
- 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 two 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 welcome the submission of original paper, review paper, survey article relevant to the all the streams of Philosophy and knowledge. The Global Journals 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 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 and Editorial Board, will become the copyright of the Global Journals.

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 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, 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 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 homepage at the judgment of the Editorial Board.

The Editorial Board and Global Journals 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 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 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.

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 (Publication Prior to Print)

The Global Journals 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 .

Informal Tips for writing a Computer Science Research Paper to increase readability and citation

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

Techniques for writing a good quality Computer Science Research Paper:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

32. Never oversimplify everything: To add material in your research paper, never go for oversimplification. This will definitely irritate the evaluator. Be more or less specific. Also too, by no means, ever use rhythmic redundancies. Contractions aren't essential and shouldn't be there used. Comparisons are as terrible as clichés. Give up ampersands and abbreviations, and so on. Remove commas, that are, not

necessary. Parenthetical words however should be together with this in commas. Understatement is all the time the complete best way to put onward earth-shaking thoughts. Give a detailed literary review.

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

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

Informal Guidelines of Research Paper Writing

Key points to remember:

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

Final Points:

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

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

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

General style:

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

To make a paper clear

· Adhere to recommended page limits

Mistakes to evade

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

In every sections of your document

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

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

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

· Align the primary line of each section

· Present your points in sound order

· Use present tense to report well accepted

· Use past tense to describe specific results

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

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

Title Page:

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

Abstract:

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

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

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

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

Approach:

- Single section, and succinct
- As a outline of job done, it is always written in past tense
- A conceptual should situate on its own, and not submit to any other part of the paper such as a form or table

- Center on shortening results - bound background information to a verdict or two, if completely necessary
- What you account in an conceptual must be regular with what you reported in the manuscript
- Exact spelling, clearness of sentences and phrases, and appropriate reporting of quantities (proper units, important statistics) are just as significant in an abstract as they are anywhere else

Introduction:

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

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

Approach:

- Use past tense except for when referring to recognized facts. After all, the manuscript will be submitted after the entire job is done.
- Sort out your thoughts; manufacture one key point with every section. If you make the four points listed above, you will need a least of four paragraphs.
- Present surroundings information only as desirable in order hold up a situation. The reviewer does not desire to read the whole thing you know about a topic.
- Shape the theory/purpose specifically - do not take a broad view.
- As always, give awareness to spelling, simplicity and correctness of sentences and phrases.

Procedures (Methods and Materials):

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

Materials:

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

Methods:

- Report the method (not particulars of each process that engaged the same methodology)

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

Approach:

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

What to keep away from

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

Results:

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

The page length of this segment is set by the sum and types of data to be reported. Carry on to be to the point, by means of statistics and tables, if suitable, to present consequences most efficiently.

You must obviously differentiate material that would usually be incorporated in a study editorial from any unprocessed data or additional appendix matter that would not be available. In fact, such matter should not be submitted at all except requested by the instructor.
Content

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

What to stay away from

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

Approach

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

Figures and tables

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

Discussion:

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

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

Approach:

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

Administration Rules Listed Before Submitting Your Research Paper to Global Journals

Please carefully note down following rules and regulation before submitting your Research Paper to Global Journals:

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

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.

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

I

K

L

M

N

O

P

Q

R

S

T

V

W