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

Global Journal of Computer Science and Technology : D
Neural and AI

DISCOVERING THOUGHTS AND INVENTING FUTURE

Highlights

- Essential Expected Computer
- Primary Education Status
- Chaotic Neural Network
- COTS Selection and Evaluation

Space Justin

Volume 12 | Issue 10 | Version 1.0 ENG

Global Journal of Computer Science and Technology: D
Neural & Artificial Intelligence

Global Journal of Computer Science and Technology: D
Neural & Artificial Intelligence

Volume 12 Issue 10 (Ver. 1.0)

Open Association of Research Society

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

Global Journals Inc.

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

Sponsors:Global Association of Research

Open Scientific Standards

Publisher's Headquarters office

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

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

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

Offset Typesetting

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

Packaging & Continental Dispatching

Global Journals, India

Find a correspondence nodal officer near you

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

eContacts

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

Pricing (Including by Air Parcel Charges):

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

Editorial Board Members (HON.)

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

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

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

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

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

Dr. Söhnke M. Bartram
Department of Accounting and FinanceLancaster University Management SchoolPh.D. (WHU Koblenz)
MBA/BBA (University of Saarbrücken)

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

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

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

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

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

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

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

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

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

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

Contents of the Volume

Global Journal of Computer Science and Technology
Neural & Artificial Intelligence
Volume 12 Issue 10 Version 1.0 Year 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Primary Education Status Analysis in Bangladesh Based on Neural Networks and Baysian Networks

By Snehasish Sarker, Md.Sarwar Kamal & Puja Das
BGC Trust University

Abstract - In this research work we have concentrate to measure the primary education status in Bangladesh, a developing country of South Asia. It known that the literacy rate of South Asian country is very slow and it is not the different in Bangladesh. Here we measure the dropout rate of primary school kids at different classes at different sessions. We have collected the data from various primary schools from Chittagong region of Bangladesh. Here we use K –Nearest Neighbor (KNN) algorithm to classify the data from irrelevant data like secondary school and tertiary level data. After then we have applied Neural Network (NN) to train the data set for better result. Finally we have compared the result by calculating the result with Bayesian Network (BN). Here we found that if the dropout rate is small Neural Network is best to measure the result and NN generate more error when the dropout rate is large. On the contrary BN is better when the rate is large.

Keywords : *K-Nearest Neighbor Algorithm, Neural Network, Bayesian Network (BN), Primary school, Dropout rate.*

GJCST-E Classification: *C.1.3*

Strictly as per the compliance and regulations of:

Primary Education Status Analysis in Bangladesh Based on Neural Networks and Baysian Networks

Snehasish Sarker [α], Md.Sarwar Kamal [σ] & Puja Das [ρ]

Abstract - In this research work we have concentrate to measure the primary education status in Bangladesh, a developing country of South Asia. It known that the literacy rate of South Asian country is very slow and it is not the different in Bangladesh. Here we measure the dropout rate of primary school kids at different classes at different sessions. We have collected the data from various primary schools from Chittagong region of Bangladesh. Here we use K –Nearest Neighbor (KNN) algorithm to classify the data from irrelevant data like secondary school and tertiary level data. After then we have applied Neural Network (NN) to train the data set for better result. Finally we have compared the result by calculating the result with Bayesian Network (BN). Here we found that if the dropout rate is small Neural Network is best to measure the result and NN generate more error when the dropout rate is large. On the contrary BN is better when the rate is large.

Keywords : K-Nearest Neighbor Algorithm, Neural Network, Bayesian Network (BN), Primary school, Dropout rate.

I. Introduction

The development of any country depend s mainly on its manpower and the pillar of good manpower is the primary education. It is easy to realize that as the children are engaging to primary education as the hopes of the prosperity will go up for respective country. Kids learning are very important for every country irrespective of rich and poor. The countries which are more developed are developed at their education level and the development start from primary level. Basically kids are very much curious in every matter and they loves to adopt with innovative culture and fashion. To ensure the kids literacy governments as well as the parents should take effective measure which will attract the kids to learn with joy and enjoyments.

Proper education for the kids can empower human beings to liberate individual mind from the curse of ignorance and darkness. It represents the foundation in the development process of any society and the key indicator of the people's progress and prosperity.

Author α ρ : Student of computer science and engineering at BGC Trust university, ID: 0913006, 0913009.
Author σ : Lecturer, computer science and engineering, BGC Trust University.
E-mail α : snehasish.bgc@gmail.com
E-mail σ : puja.cse.bgc@gmail.com
E-mail ρ : sarwar.saubdcoxbazar@gmail.com

In the view of the importance of education to a country like Bangladesh the present thesis addresses limitations of primary education system, which is diversified and multifarious due to economic, socio cultural, political, regional and religious factors. The entrance of primary education is maintained mainly by the government of Bangladesh. More than 75% schools are controlled by the government and around 83% of the total children enrolled in the primary level educational institution go to these schools (Baseline Survey, 2005:3). Similarly, more than 70% primary teachers are working in the government controlled schools. Besides government run primary schools, nine other category of primary schools are administered, monitored and maintained by different authorities. Disparity and lack of coordination among these institutions constrains the attainment of universal primary education and in its effort to increase enrollments and quality education. Variations in teacher student ratio, the number of qualified and trained teachers between the categories, also pose a big challenge towards achieving the goals of universal primary education.

In the circumstances of the open scenario, Bangladesh became one of the signatories to the UN Millennium Declaration in 2000, and has achieved to eight Millennium Development Goals (MGDs) that affirmed a perception for the 21st century (Burns et al, 2003:23). Bangladesh also pledged to implement the MDGs roadmap by 2015. The MDG-2 targets for 'Achieving Universal Primary Education' are claimed to be on track in Bangladesh, showing remarkable achievements in terms of net enrolment rate in primary education 73.7% in 1992 to 87% in 2005 and primary education completion 42.5% in 1992 to 83.3% in 2004 (Titumir, 2005:120). Bangladesh government itself had taken many initiatives, including the Compulsory Primary Education Act 1993, which made the five-year primary education program free in all primary school. The government adopted demand side intervention policies such as food for education program and stipend program for primary education. Of late, the government introduced primary education development program (PEDP-II), a six year program beginning in the year 2000, which aims to increase access, quality and efficiency across the board in the primary education sector. Despite existing socioeconomic problems,

Bangladesh by now has achieved a good progress in net enrollment rate and education completion rate in primary education. The current paper will examine the outcomes and challenges that have emerged as a result of Bangladesh being the signatory of MDG-2 for achieving universal primary education by 2015. It would further investigate whether the target of the second Millennium Development Goal is attainable within the stipulated time.

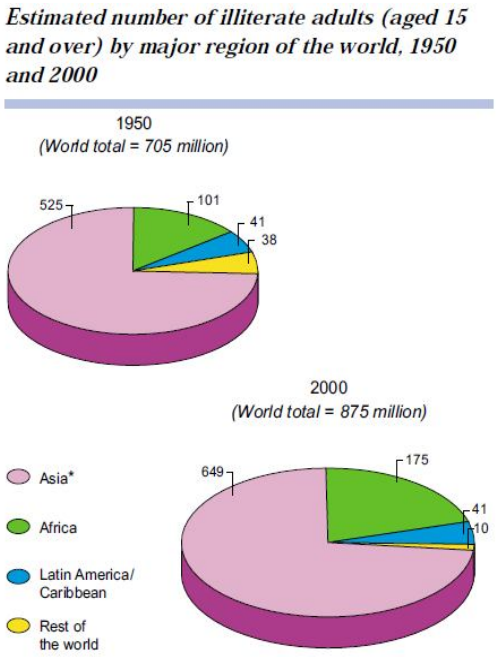
To better improvement of literacy at primary level we examined and interviewed a lot of kids at various places and found that the children are used to learn with joy and entrainment. Recently in Bangladesh various primary school at North Bengal part of Bangladesh implemented park at primary school compartment and get the very positive results. Hathazari Upazila parishad has arranged an innovative approach to help bring down dropouts in government primary schools. It is setting up children's parks in schools and constructing roads to give students easy access to schools. Encouraged by cent percent pass rate in last year's class-V terminal examinations in the upazila in Patuakhali, the parishad set up children's parks in 60 of 122 primary schools [3] there to draw students. Construction of 50 more is underway. It also paved 25 roads for easy access to schools for [3] students from nearby areas. About 65 more are under construction. Urmee, a student of class-IV at Madanpura Government Primary School said she could not attend school in the rainy season last year for bad condition of the road to her school. But she has not missed any classes this year after the road got repaired. Miss Dhunu Chokrabarty headmaster of the Mujaffarabad govt primary school said most students now come to school in the rainy season but the attendance hovered around 70-75 percent during the same period before the road was repaired. The children's park in the school is an additional attraction for students, he added. Several students were found waiting for their turn to get on a swing at Hosnabad Government Primary School. Many of them come to school early to play on the swings. Rafiqul Islam, a student of class-IV of Najirpur Government Primary School, said their playground used to go under water in the rainy season. It was developed last year, and now they can use it all year round Dilip Chandra Sarker, headmaster of Baromashia govt primary school, said the parish ad's initiative helped increase attendance rate in his school. Now about 93 to 97 percent students attend classes, while it was 80 to 85 percent a year ago. Ayesha Akter Chowdhury,[3] headmaster of Sholkata govt primary school, said the children's park of the school has brought a big change. The attendance rate went up to more than 90 percent from below 80 percent. Upazila Chairman Engineer Mojibur Rahman said they are setting up children's park on school playgrounds to provide students with leisure facilities on instructions from Patuakhali-2 lawmaker ASM Feroz, also whip of the Jatiya Sangsad. The parishad is spending about Tk 50,000 on each children park from its own fund. The project started in January 2010 and will end by December next year."We have already re-excavated 36 school ponds. Of them, 10 have been brought under fish cultivation with fisheries department's help. Another 63 will be re-excavated by next year," said Mojibur, who last year received an award from the prime minister for his role in promoting primary education in his area. "We will take another programme to grow seasonal vegetables on unutilised school land. The profits from the sale would go to poor and helpless students for buying books and stationery," he said. The parishad published a 472-page book Tathya Kanika. It contains names of all educational institutions in the Upazila and information on students and teachers."We have already distributed the book to all educational institutions in the area. It will help students get in touch with each other. Upazila officials will also be able to communicate with any teachers or institutions by using the information provided in the book," said Mojibur.Md Ibrahim, Upazila primary education officer, said "We are getting good results from these steps."Now most students go to school that has become a more interesting place for them, he added.

The organizations of the work start by the literature overview after the introductory descriptions. At the introductory description we have observed the situation of Bangladesh primary education and the current status of the country. At the literature study part we have look forward towards the whole world situations specially the developing courtiers of the world like Ethiopia, Sudan, Nigeria and Nepal. The UNESCO report also helps us a lot towards the exact scenario of the education status of the developing country. The data collection is done after the literature study and the data set are real world data from various primary school of Chittagong region of Bangladesh. It is very alarming that we have found very much irregularity to collect the data. Data collection helps us to design the intelligent system for our desired work. At first we applied the data classification techniques to the collected data. We choose K-Nearest Neighbors (KNN) algorithm for data set classification. The data set classification helps us to reduce the redundant data. At last but not the least we compare the result by both Neural Networks (NN) and Bayesian Networks (BN).

II. Literature Over View

We have studied and checked related supporting documents and information towards the education status analysis throughout the world irrespective of rich and poor country. But in this work we have concentrate to design and investigate the condition of developing country. We found that there are some paper support the qualitative measurements and

some other had concentrate on survey based measurements. From the study of Ethiopia, one of the backward countries at education and economic condition we see that only survey based analysis is done [1]. Data are collected from students, instructors, gender officer and from guardian though the interview, questionnaires and discussion with focus group. Here no advanced technique is used like Neural Network (NN) and Bayesian Network (BN) or data mining techniques as data pre-processing, data warehousing, linear regression, non linear regression and so on. But for the generic evaluation and measurement it is very essential to have to have an Intelligent Decision System (IDS) at every levels of education. According to the World Education Report 2000 we have noticed that at South Asia there is huge part of the children are out of education and they are suffering a lot of socio-economic problems like poverty, proliferating acts of violence and conflicts, illiteracy and rich and poor gap. The survey by the expert of UNESCO found and 649 [2] million illiterate resides at south Asian region. The figure 1 bellow depicts the matter.



* For 1950, data for the Asian countries of the former USSR are included in Rest of the world.
Source: Figures for 1950 are taken from World Illiteracy at Mid-Century, p. 15, Paris, UNESCO, 1957.

Fig. 1 : The illiteracy comparison at 1950 and 2000

All the reports of UNESCO and survey paper of the researcher are based on people interaction and qualitative. But run the quick and accurate result it is convenient to have an intelligent system. If we want to perform the calculation to the Robot what will be situation than? We must need the Decision Support System (DSS).

III. Data Collection

We have collected the data set from the Chittagong region of Bangladesh for the experiment of our research work which is done by the assistance of continues support by the head teacher of the schools. We mainly concentrate data set for the classes where the students enrolled at first class and those are appeared for the Primary School Certificate (P.S.C) exam. In this condition we found that the measurement is efficient and meaningful for the evaluation as well as for the decision. Our sample spaces were the Mujaffarabad Government Primary School, Hathazari, Baromashia Government Primary School, Fatikchari, Sholkata Government Primary School, and Anwara. From Mujaffarabad Government Primary School, Hathazari we have collected consecutive data from 2004 to 2008 for each year from class one to class five. The table 1 to 5 narrates the collected data for various batches from 2004 to 2012. For table 1 we collect the data set for class one to class five for a respective batch. Here we can see that from 2004 to 2005 four students have dropout among thirty students. At 2006 we see that there are more students than class one due to the dropout of the previous class those who are enrolled at class one at 2003. For a certain batch of 2004 at class one. We get some variations at class three, four and five due to the reasons of dropout of the previous classes.

Year	Class	Total Students
2004	Class 1	30
2005	Class 2	26
2006	Class 3	35
2007	Class 4	32
2008	Class 5	33

Table 1 : Mujaffarabad Government Primary School, Hathazari

As the same table 2 depicts the data set for the same school for another batch that is started from 2005 for class one and end the class five at 2009.

Year	Class	Total Students
2005	Class 1	37
2006	Class 2	33
2007	Class 3	40
2008	Class 4	46
2009	Class 5	36

Table 2 : Mujaffarabad Government Primary School, Hathazari

This table also indicates the dropout rate from class three to class five and some new students also get enrolled from class three to class five. It is really pathetic to see that at village part the kids are out of school at various ages due to the lack of information or other socio-economic problems. The both tables above indicate the data set from rural part of the Bangladesh.

The table 3 below also shows the same data set for the same school but the timeline is different than other two tables. Here we see that the enrollment rate increase over the time. It is also positive side for Bangladesh that the students and the guardians are now becoming causes than earlier.

Year	Class	Total Students
2006	Class 1	45
2007	Class 2	38
2008	Class 3	32
2009	Class 4	29
2010	Class 5	26

Table 3 : Mujaffarabad Government Primary School, Hathazari

We than collect the data set from the area where the people are more causes than the previous area and we have observed a interesting change towards the developments of the literacy. At this place the government involvements are also very frequent than the previous area.

Year	Class	Total Students
2007	Class 1	129
2008	Class 2	119
2009	Class 3	112
2010	Class 4	98
2011	Class 5	85

Table 4 : Baromashia Government Primary School, Fatikchari

The table 4 contains the data more than other three tables and the dropout is also more than the others. If the data size is more the dropout is also more. The main reasons behind the dropout is that the poverty and guardians inconsistency for study and literacy. Besides they have the idea what will happen after study, for job they needs lobbing. Here for Fatikchari area we have found that majority of the people are living at middle east for earning as result the students enrollment rate increase for the better financial support. Similarly at table 5 the enrollment is smaller than the Fatikchari area due to the same facts.

Year	Class	Total Students
2006	Class 1	65
2007	Class 2	57
2008	Class 3	50
2009	Class 4	48
2010	Class 5	38

Table 5 : Sholkata Government Primary School, Anwara

IV. K-Nearest Neighbor (K-NN)

K–nearest neighbor (K-nn) algorithm is a branch of supervised learning. Now-a-days it is being applying in various fields of data and information processing irrespective of science, commerce and arts. In the context of machine learning, K-nn is considered an effective data classification technique based on adjacent developed examples of sample space. The value of K is always positive and an object is classified by considering the greater number of choice of its neighbors. The neighbors are chosen from data set which is best fit for correct classifications and Euclidean distance helps to measure the overall distances. Here every occurrence correlates to points in sample space or within populations. Generally distance or similarity between instances or objects is easy if the data sets are numeric or integer. A very typical formula to calculate distances is Euclidian distances formula as follows:

$$d = \sqrt{((x_{1i}-x_{1j})2+ (x_{2i}-x_{2j})2)} \quad (a)$$

In some cases Manhattan or City Block distance also applicable:

$$d = (x1i-x1j) + (x2i-x2j) \quad (b)$$

However it is very essential to bear in mind that all the instances at sample space must be same scale. As for example income will compare with income not the height of the human beings.

For qualitative data the distance measurement process will be different and it is important to consider that the instances are same or not. At this stage the qualitative objects are measured by allocating Boolean values to each object. It might be possible to converts to instances between which distance can be identified by some techniques. As for example color, temperature, age, height etc. Text and character has identified as one instance per word with the frequency start from 0, 1, 2……n.

a) The classifications process of K-nn as follows

The two main steps of K-nn must follow are:

1. Training
2. Predictions

Training means to get information from all sample spaces and populations. To accomplish this work we need to have the idea about the all instances and objects. In this sense it is very important to bear in mind that data set must be in same class. The qualitative and quantitative data measurement will be different.

The predictions will manage by considering the predefined methods.

b) The k-nn Algorithm

The total algorithmic steps are as follows:

1. Parameter selections (int m, int n). m=0, n=1, 2, 3……n.
2. Distance calculation

   $\sqrt{\Sigma}$ (q-pi) 2 where i=0,1,2,3……..,n

3. Short the distances of sample space and marked the closet neighbors in the context of K-th smallest distance.

SHORT NEIGHBORS (**S**, **C**)
Input instances **S** with n sample objects, comparator **C**
Output instances S sorted according to **C**
if S.length() > 1
Then (S_1, S_2)←**divide**(S, n/2)
SHORT NEIGHBORS (S_1, C)
SHORT NEIGHBORS (S_2, C)
S ← SHORT NEIGHBORS (S_1, S_2)

4. Similarities assumption :Instances that are close together should have similar values Minimize

$$\xi(f) = \sum w_{ij}(f_i - f_j)^2$$

Where w_{ij} is the similarity between examples i and j.
And f_i and f_j are the predictions for example i and j.

5. Predict the value as follows:

Standard KNN $\hat{y} = \arg\max_y C(y, Neighbors(x))$

$$C(y, D') \equiv |\{(x', y') \in D' : y' = y\}|$$

Distance-weighted KNN

$$\hat{y} = \arg\max_y C(y, Neighbors(x))$$

$$C(y, D') \equiv \sum_{\{(x', y') \in D' : y' = y\}} (SIM(x, x'))$$

$$SIM(x, x') \equiv 1 - \Delta(x, x')$$

6. Find out the best heuristics distance

```
f(n) = g(n) + h(n)
```

Where:

- g(n) is the cost of the best path found so far to n
- h(n) is an admissible heuristic
- f(n) is the estimated cost of cheapest solution through n

V. New Maximum Nearest Area (NMNA)

How k-nn selects the desired values from a lot of alternatives is that it calculates its nearest most predicted value. The following figure depicts the computations.

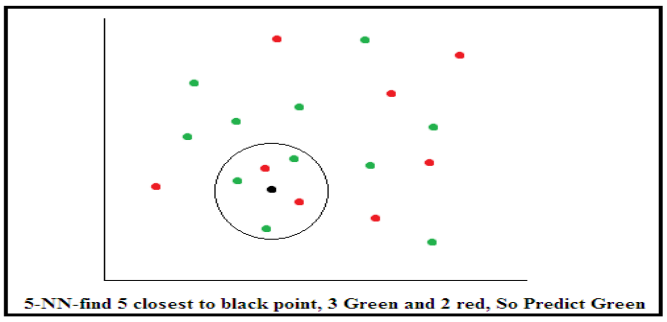


Fig. 2 : The neighbor selecting process

In the figure above we see that the small circle belongs three different color dots where the black one is the pivotal element and based on that point we will calculate the green and other two green and red points. According to this figure we have to predict the green points as a K nearest neighbors. The neighbors are very closest to the pivotal point. It is vital point that New Maximum Nearest Area (NMNA) must be accurate otherwise it will not work properly for data set selections. In the k-nearest neighbour process, the only benefit of selecting a large k value is to scale down the variance of the conditional probability estimate. By parallel, in oursystem, a large k value can probably lead to a large confidence measure, therefore to a small probability of error. It is very much essential for each inquiry point of the confidence measure and the corresponding probability of error can be easily computed for different numbers of neighbors. Therefore, one can choose to increase the neighborhood size k until a preset probability of error threshold is achieved. So, the probability of error, or equivalently the confidence level, provides a mechanism to dynamically determine the size of the neighborhood. We will call the modified version of the k-nearest-neighbor rule the confident-nearest neighbor rule.

VI. Organization of the Process

Now it is important to build the process how K-nn may organized in reality or the time line. To manage the proper training area we have to shorten the area or to select the appropriate area. When we are able to fix the sample area for computation, it will help us to reduce the computational complexity for entire process. To accomplish the total work we must follow some preconditions. First we have to choose a population area where we can apply our algorithm to extract our outcome. The figure 2 bellow shows a population area for our research activity.

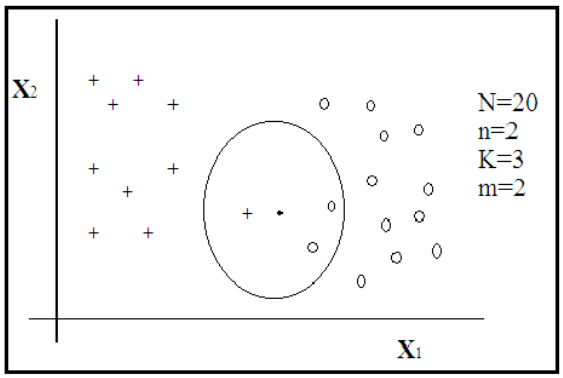


Fig. 3 : The total population area for K nearest neighbors

Here,
N = Total number of data set at population space. In this figure above we see that there is twenty (20) objects are outside the circle. The circle denotes the selected sample space. Inside the circle the black point indicate the pivotal or central point.

K=the total neighbors. Here the value of K is three (3).
n= indicate the nearest value.
m= categories of the neighbors. In the figure above
We see that there are two categories of neighbors.
One data set indicate by plus (+) sign and other is small hole.

At the beginning we narrow the area as sample space from population area. The figure 3 below shows the desired sample space for our working activity. By the reduction process, we are able to reduce forty percent of the computational cost.

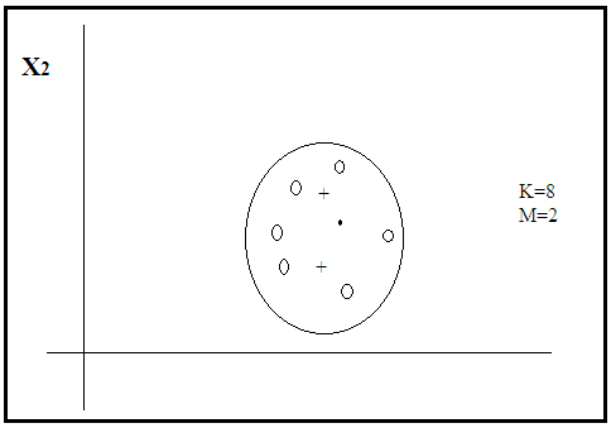


Fig. 4 : The reduction sample area for K nearest neighbors area

By comparing both figure 2 and figure 3 we get the followings equation as a general format.

$$\text{Computational reduction cost (CRC)} = \frac{Samplespaceneighbors}{Populationneighbors}.$$

Suppose at figure 2 we see that there are twenty objects or data set and after making classification based on K nearest neighbors algorithmic techniques we have extracted as figure 2 where there is only eight data set or object which are very much closed to pivot point indicated inside the circle as black point.

$$\text{Here the CRC} = \frac{08}{20} \times 100$$

$$= 40\ \%$$

It is really interesting and effective that for twenty data points we have reduced 40% of the computational cost and as a consequence it reduced the complexity of the computation. From the time and space complexity view point in linear search we get the better improvement at both time and space complexity. Time complexity defines how the algorithm behaves while the input size increases. Generally in the worst case, the running time is proportional to n where n denotes as input size. For any input size with time complexity O(n) will take the twice time while the input size doubles as a consequence with time complexity O(N2) will take four times longer while input doubles and so on. As a average case the time complexity is as follows:

C (n) =1*1/n+2*1/n+3*1/n+.....+n*1/n = (n+1)/2.
So here we get the improvements of (n+1)/2 * 40 %
=(n+1)/2.*40/100
=40* (n+1)/200.
= (n+1)/5.

Hence we see that the CRC helps to cut the cost to one fifth of the original cost. So the time complexity is the CRC linear search is O (n/5) where the total population cost is O (n) for the n input size. It will be easy to remember that the result of comparison will change when the input size will change. From further reduction we get the figure below at figure 4. According to the equation a we manage our desired sample data set configured as bellow.

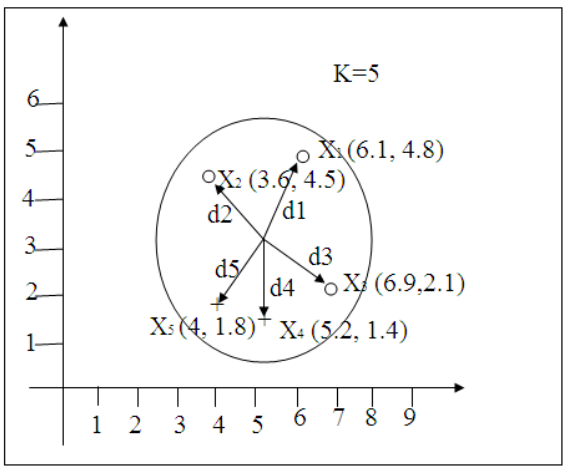


Figure 5 : The final reduction of the K nearest neighbors

At figure five we depict the problem domain by a time line diagram. The time line diagram behaves the similarities of Support Vector Machines (SVM) of Maximum Margin Hyper-plane (MMH). The pivot value is selected as the mean point of the data set.

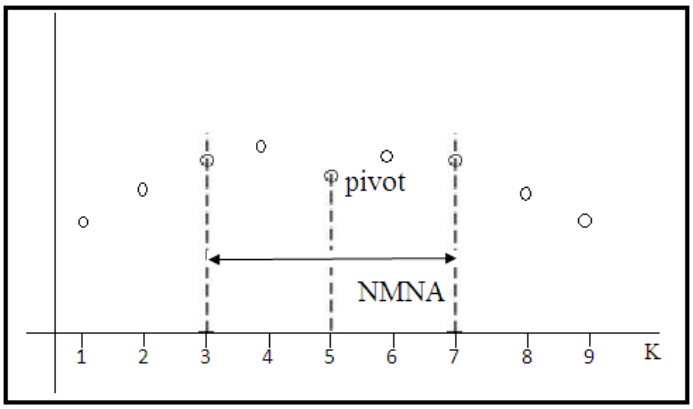


Figure 6 : The time diagram of K nearest neighbours as a MMH of SVM

VII. Neural Networks Perception

Neural networks represent a brain analogy for information processing. These models are biologically exhilarated rather than clear-cut clone of how the brain actually functions. The figure 7 shows the similarities between artificial neural network and biological neurons. Neural ideas are usually implemented as system simulations of the massively parallel processes that involve processing elements interconnected in network architecture.

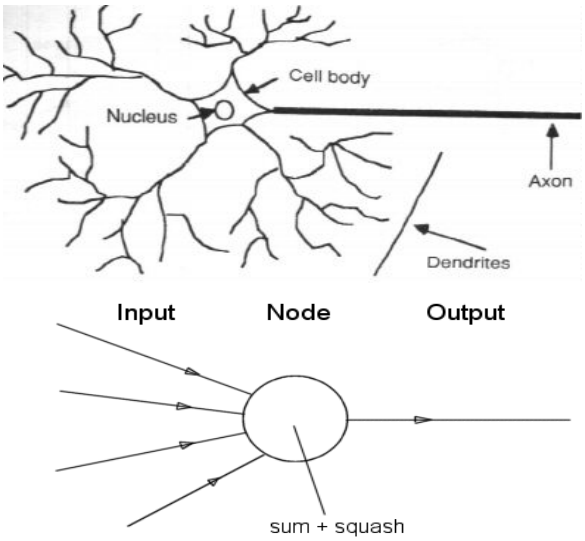


Fig. 7 : The biological and artificial neurons

Neurons receive the sum of information from other neurons or the external elements, perform transformation on the inputs and then pass the transformed information to other neurons or the external outputs. A typical structure is shown in Figure 8. For the better measurement and accurate result we have experimented by Multi Layer Involvement (MLI) of the Neural Networks (NNs) which is an advanced computational and learning method at modern computation and Intelligent Systems (ISs). A MLI consists of three layers named input layer, hidden layer, and output layer. A hidden layer receives input from the previous layer and converts those inputs into outputs for further processing. Several hidden layers can be placed between the input and output layers, although it is quite common to use only one hidden layer. Every working cell is connected with each of other cell as directed graph. A NN is very much similar with a directed graph where the neuron cells are considered as vertices and the connections between the cells are edges. In that case each edge is associated with its weights and the weights must reflect the measurements of the input and output results. Naturally an Artificial Neural Network (ANN) is consisting of Adaptive Linear Neural Elements (ADLINE) that changes its structure according to the propagation of information on external or internal matters through the network during the learning phase.

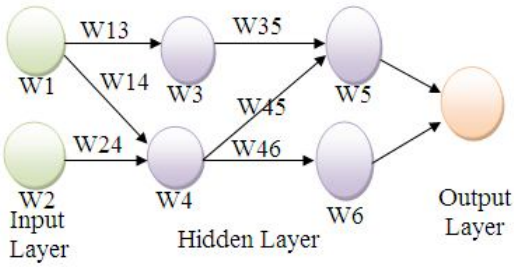


Fig. 8 : Multi layer neural network

Neural Networks learn to adapt the inputs to produce desired outputs. When the NNs choose to learn by supervised process than the learning must be inductive. To learn the NN first compute the temporary output, then compare the output with target output and finally adjust the weight and repeat the process. When existing output are available for comparison, the NN start the learning process. The figure 9 bellow shows the learning process of a single neuron. The weight adjustment with learning is $\Delta W_i = \eta * (D-Y).I_i$.

Where,

η=learning rate, D= Expected output, Y= actual output, I= input.

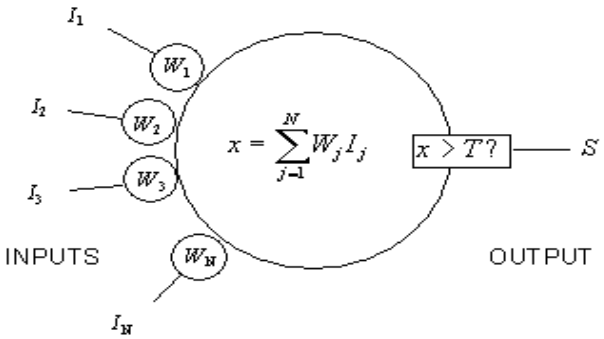


Fig. 9 : The learning process for a single artificial neuron

For a Multi Layer Neural networks the mathematical learning process is as follows:

$$y_k^1 = \frac{1}{1+e^{-w^{1kT}x-a_k^1}}, k = 1,2,3 \quad \text{(c)}$$

$$y^1 = (y_1^1, y_2^1, y_3^1)^T$$

$$y_k^2 = \frac{1}{1+e^{-w^{2kT}y^1-a_k^2}}, k = 1,2$$

$$y^2 = (y_1^2, y_2^2)^T$$

$$y_{out} = \sum_{k=1}^{2} w_k^3 y_k^2 = w^{3T} y^2 \quad \text{(d)}$$

For error measurements the equation is

$$E = \frac{1}{N}\sum_{t=1}^{N}(F(x_t;W) - y_t)^2 \qquad (e)$$

To change the weight the learning equation are

$$\Delta w_i^j = -c \cdot \frac{\partial E}{\partial w_i^j}(W) \qquad (f)$$

$$w_i^{j,new} = w_i^j + \Delta w_i^j$$

C is a learning parameter. Usually it is constant.

VIII. Bayes'Theorem

Bayes' theorem and conditional probability are opposite to each other. Given two dependent events A and B. The conditional probability of P (A and B) or P (B/A) will be P (A and B)/P (A). Related to this formula a rule is developed by the English Presbyterian minister Thomas Bayes (1702-61).According to the Bayes rule it is possible to determine the various probabilities of the first event given the outcome of the second event in a sequence of two events.

The conditional probability:

$$\text{P(B/A)} = \frac{P(AandB)}{P(A)} \qquad (1)$$

The equation (1) will help to find out the probabilities of B after being occurrences of the A. we get the Bayes' theorem for these two events as follows:

$$\text{P (A/B)} = \frac{P(A).P(B/A)}{P(B)} \qquad (2)$$

If there are more events like A1, A2, and B1, B2.In this case the Bayes theorem to determine the probability of A1 based on B1will be as follows:

P(A1/B1)=

$$\frac{P(A1).P(B1/A1)}{P(A1).P(B1/A1) + P(A2).P(B2/A2)}$$

IX. Implementation

The implantation is done according to the concepts and knowledge of NNs and Bayesian Networks. The following figure 11 narrates the brief descriptions of the implementation. The fundamental algorithmic steps perform as the core idea or engine for the calculation. Our proposed algorithm for dropout prediction is as follows:

The algorithmic steps (Proposed)

1. Start
2. Take node number n.
3. Find out the probability P(A) for all nodes.
   P(A)=((A-b)/A)*100;
4. Find out the average probability for all nodes.
   Average probability = (P(A)+P(B)+…….+P(N))/n;
5. Let the average probability is equal to the probability of last node. P(Y)= Average probability of all nodes
6. Find out the result node X.
   X=Y-((P(Y)*Y)/100);
7. End
8. Repeat steps 2 to steps 6 again.

To accomplish the full work with efficiency and accurate we have worked firstly on input data sets and fit them to the networks. In this work we set input as a numeric data for the Neural Network. Here four inputs are applied to the network to predict the output of the class five. Our processing shows in figure 11.

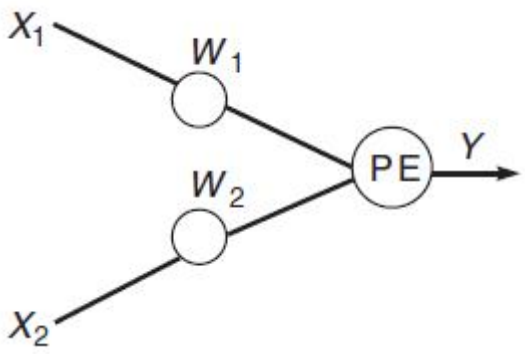


Fig. 10 : A processing element with two inputs set

X1 and X2 may be the numeric value or representations of an attribute. W1 and W2 are the weights for the networks and Y denotes the output of the inputs. PE indicates the Processing Element. The simple calculation for the output Y = X1 W 1 + X2 W 2. The output indicates the salutation of the inputs. At the output level the ANN provides the value 1 for yes and 0 for no. another important part for the processing is weights of the network. Weights are the relative strength of the inputs set or many weights transfer input from layer to layer.

```
$w=((($a-$b)/$a)*100);
$x=((($b-$c)/$b)*100);
$y=((($c-$d)/$c)*100);
$z=(($w+$x+$y)/3);
$e=((($z*$d)/100)+$d);
echo "According to Newral network the result is ".$e;
```

Fig. 11 : The brief idea for NN implementation

One of the most important functionality of this work is the error learning training to get the desired outputs. The figure 12 shows the first pass of our research work. Gradient of the neuron= **G** =slope of the transfer function×[Σ{(weight of the neuron to the next neuron) × (output of the neuron)}].

G1=(0.7265)(1-0.7265)(0.0397)(0.5)(2)=0.0093.
G2=(0.6508)(1-0.6508)(0.3492)(0.5)=0.0397.
G3=(1)(0.4292)=0.4292. Error =1-0.4292=0.5708.

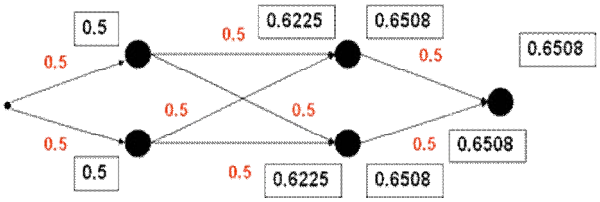


Fig. 12 : Computation at NN for the collected data set

To minimize the errors rate we have adjusted the weights which have shown bellow at figure 13.the new weights is New Weight=Old Weight + {(learning rate)(gradient)(prior output)}.
0.5+(0.5)(0.0093)(1),0.5+(0.5)(0.0397)(0.6225,
0.5+(0.5)(0.3492)(0.6508).

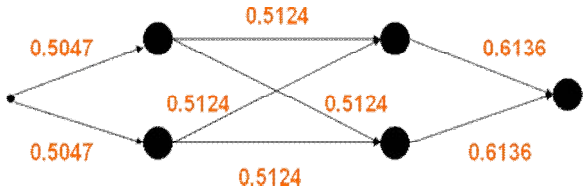


Figure 13 : The weight adjustments

Finally we have implemented for Bayesian network and we founds as figure 14. According to the operation of the Bayesian networks and application in this work we have noticed that BN runs as follows.

```
$total=($a+$b+$c+$d);
$i=($b/$total);
$j=($c/$total);
$k=($d/$total);
$l=(($i+$j+$k)/3);
$r=($l*$total);
echo "According to Baye's theuram the result is ".$r;
```

Figure 14 : The Bayesian implementations

X. Result

The results of this implementation snaps are shows bellow.

Input students number in the text field

Class 1: 129
Class 2: 119
Class 3: 112
Class 4: 98
Submit

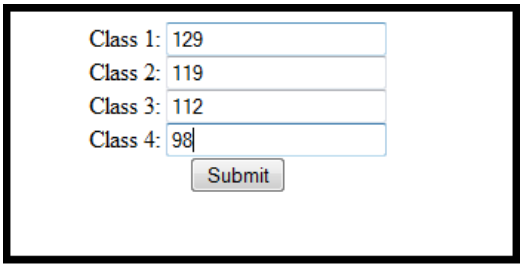


Figure 15 (a) : The inputs for the Baromashia School

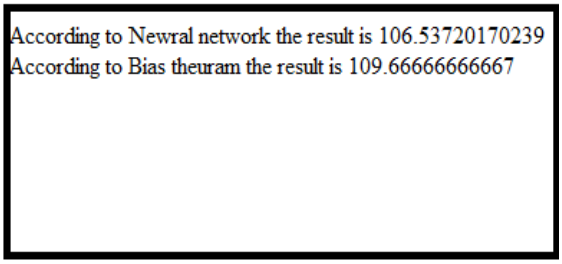


Figure 15(b) : The outputs of the previous inputs

Input students number in the text field

Class 1: 30
Class 2: 26
Class 3: 35
Class 4: 32
Submit

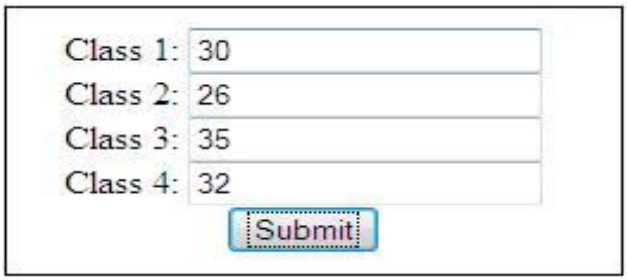


Figure 16 (a) : Inputs for Mujaffarabad School

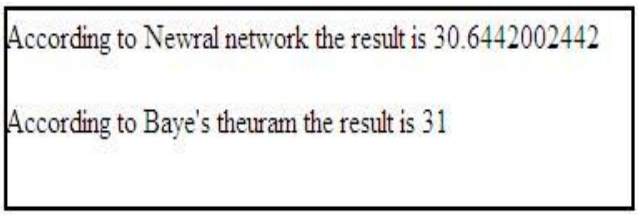


Figure 16 (b) : The outputs of the previous inputs

XI. Comparisons

The result of the both NN and BN are very effective for this research. Both procedure are very suitable for machine to learn and predicts the accuracy of the result. Machine learning is very essential for current era. In the age of information superhighway every steps of computations are becoming machine oriented and are being handing based on automated way. Though the efficiency we make a comparison and found that NN is better at the cases when the input size has the less difference in dropout rate and on the contrary the BN is better while the input dropout rate is larger.

NN is 96% is corrects and BN is 91.5% accurate in this research.

XII. Acknowlefgement

We are thanking the head masters of the schools who help us by providing the data regarding the enrollments and dropout students from their schools. We are showing our respect to Miss Dhunu Chokrabarty sir, head teacher of Mujaffarabad govt primary school, Mr. Dilip Chandra Sarker sir head teacher of Baromashia govt primary school, and Miss. Ayesha Akter Chowdhury head mistress of Sholkata govt primary school for their regular support for data collection and analysis.

XIII. Conclusion

It is very good to say that this research will help to assess the degree of dropout students from any country especially from the developing country. We have measured the significant amount of dropout students from primary stage due to the various socioeconomic problems like lack of knowledge, poverty, and social barrier. Our implementation is very efficient for automated system as well as machine learning. Besides this work we noticed a few drawback regarding the time line and data collections and organization. At future we will overcome the problems regarding the indicated problems.

References Références Referencias

1. Trend and causes of female student's dropout from teacher education Institutions of Ethiopia: the case of Jimma university by Wudu Melese, Getahun Fenta. Ethiop. J. Educ. & Sc. Vol. 5 No. 1 September, 2009.
2. World education report 2000 by UNESCO.
3. The daily star Bangladesh.
4. NEURAL NETWORKS IN DATA MINING by DR. YASHPAL SINGH, ALOK SINGH CHAUHAN, Journal of Theoretical and Applied Information Technology.
5. Effective Data Mining Using Neural Networks, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 8, NO. 6, DECEMBER 1996 by Hongjun Lu, Member, IEEE Computer Society, Rudy Setiono, and Huan Liu.
6. Evolving Artificial Neural Networks, by XIN YAO, PROCEEDINGS OF THE IEEE, VOL. 87, NO. 9, SEPTEMBER 1999.
7. Principal Feature Classification by Qi Li and Donald W. Tufts, IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 8, NO. 1, JANUARY 1997.
8. Neighborhood size selection in the k-nearest neighbor rule using statistical confidence, JigangWang, Predrag Neskovic, Leon N. Cooper.
9. Tom Mitchell, Machine Learning, McGraw-Hill Computer science series, 1997.

Global Journal of Computer Science and Technology
Neural & Artificial Intelligence
Volume 12 Issue 10 Version 1.0 Year 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Artificial Intelligence an Essential Expected Computer World Surveillance

By K.Sudheer Kumar, D.Srikar & CH.S.V.V.S.N Murthy
D.N.R College, Bhimavaram

Abstract - A position paper toward an important and urgent discussion on how best uses the potential of Artificial Intelligence in the context of Computer World surveillance. AI is often cited in papers on Computer World surveillance. But what is meant is using pre-existing AI techniques in Computer World surveillance. AI techniques are established around applications. Computer World surveillance has never been an area of deliberation in AI. In this paper we argue that Computer World surveillance calls for new and specific AI techniques developed with that kind of application in mind. In practice, this paper is based on a broad overview of different slants, which have the budding to be game changers in Computer World surveillance. This paper focuses on web solicitation security and supporters the use of Knowledge Based Systems, probabilistic reasoning and Bayesian apprising to control the probability of false positives and false denials.

Keywords : *Documentation, Security, Theory, Bayesian updating, CSRF, probabilistic reasoning.*

GJCST-E Classification: *I.2.1*

Strictly as per the compliance and regulations of:

Artificial Intelligence an Essential Expected Computer World Surveillance

K.Sudheer Kumar [α], D.Srikar [σ] & CH.S.V.V.S.N Murthy [ρ]

Abstract - A position paper toward an important and urgent discussion on how best uses the potential of Artificial Intelligence in the context of Computer World surveillance. AI is often cited in papers on Computer World surveillance. But what is meant is using pre-existing AI techniques in Computer World surveillance. AI techniques are established around applications. Computer World surveillance has never been an area of deliberation in AI. In this paper we argue that Computer World surveillance calls for new and specific AI techniques developed with that kind of application in mind. In practice, this paper is based on a broad overview of different slants, which have the budding to be game changers in Computer World surveillance. This paper focuses on web solicitation security and supporters the use of Knowledge Based Systems, probabilistic reasoning and Bayesian apprising to control the probability of false positives and false denials.

GeneralTerms : Documentation, Security, Theory.
Keywords : Bayesian updating, CSRF, probabilistic reasoning.

I. Introduction

It has been known for a long time that in Computer World surveillance, defense uses a flawed hypothesis because it leads to a strategy based on tweaking and "fixing the comprehending" with no long term vision. When new forms of attacks appear, ad hoc response are put together, which often involves making the use of the internet more cumbrous, by adding layers of authentication. "Defensive measures tend to involve complicating protocols or their implementation, making things more secure by making them more cumbersome a mentioning abstaining from using some functionality, as more often than not each new functionality provides new points of entry for malicious activities [1],[2],[3],[4]. But the introductions of new functionalities are precisely what make the internet so attractive and successful. Furthermore there is still no tolerable defence to zero day attack, as anomaly based detection has still some open problems, like the false positive probability. An ideal cyber- defense would provide full protection to users, while preserving all the functionalities. We are very far from this situation. But there is no reason why in the long run, we could not get close to such a situation.

One thing that cyber-defense can do and should is to be more intelligent. The approach to defense based on" fixing the plumbing" is inherently suboptimal. A massive paradigm shift is needed, the kind of paradigm shift that makes a much heavier use of Artificial Intelligence (AI). The idea or making heavier use of AI in Computer World surveillance is not new. In an editorial in IEEE security and Privacy [5], Carl Landwehr stated that "In their early days, computer security and artificial intelligence didn't seem to have much to say to each other. AI researchers were interested in making computers do things that only humans had been able to do, while security researchers aimed to fix the leaks in the plumbing of the computing infrastructure or design infrastructures they deemed leak proof." But the dream of retroactively make the internet secure and leak-proof, is clearly not a clever approach.

The introduction of new technologies like the proliferation of new web applications or the increasing use of wireless, have exacerbated this fact [1]. Computer World surveillance, has become the most complex threat to society. Despite years of incremental improvements in cyber-defense, it is clear that a paradigm shift is needed, but at the same time difficult to imagine.

This is even more true for web-application security and the need for AI is even more obvious and urgent there. Against web application attacks, such as Cross Site Scripting (XSS), Cross Site Request Forgery (CSRF), injection code, the present approach consists in introducing rules supposed to prevent them. This is the same kind of logic that un derived the idea of same origin policy. Over the years XSS and CSRF began to mean a variety of attacks. Some of them can be construed as direct circumvention of the same origin policy. Same origin policy looked like a simple and efficient protection.

It turned out that it could be circumvented reasonably easily and was preventing some functionality to modern websites. In the words of D. Crockford same origin policy (which is adopted by most browsers) "prevents useful things and allows dangerous ones" [7]). Today, this policy is being revisited. For example, cross-site access control [8] is an attempt to refine the security policy in such a way that one can get the benefit of cross site access without the security implications.

Author α : Asst.Professor, Department of Computer Applications, P.G.Courses & Research Center, D.N.R College, Bhimavaram.
E-mail : sudheerkumarkotha@gmail.com
Author σ : Associate Professor Department of Computer Application P.G.Courses K.G.R.L.College, Bhimavaram.
E-mail : srikar_d@yahoo.com
Author ρ : Asst. Professor, Department of IT Sri sai Aditya Institute of Science &Techonlogy, Surampalem.
E-mail : chsatyamurthy@gmail.com

In the same way, SQL injection codes were described as "extremely easy to avoid" [9]. Still recently MYSQL and Oracle, both suffered such attacks. What may be construed as mistakes is also reflection of the fact that those mistakes are not so easy to avoid in the increasingly complicated world of web applications. To detect web application attacks such as XSS, CSRF or injection codes requires more than simple rules, but the ability to some form of context dependent reasoning. Despite some work done in the past, AI does not play central role in Computer World surveillance today and Computer World surveillance has not been an area of development of AI as intensely pursued as robots, machine learning and the like. Typically, the use of AI in Computer World surveillance has consisted in using some tools developed in AI and apply them to intrusion detection or other aspects of Computer World surveillance. The approach consisting in importing some AI techniques developed in totally different areas and try to apply them to Computer World surveillance, may work in a few cases, but has inherent and severe limitations. AI has been developed and is organized around specific applications, many of them (AI is a large field). Computer World surveillance has specific needs and to be a long term contributor to Computer World surveillance, new AI techniques will have to be developed specifically.

Obviously AI has made a lot of accomplishments and there is a lot to be learned relevant for Computer World surveillance and many new techniques suited for Computer World surveillance could be inspired from existing ones in AI. In a sense this has happened already. As observed by C. Landwehr [5]: "A branch of AI that has been connected with computer security from relatively early days is automated reasoning, particularly as applied to programs and systems. [...] Although it wasn't identified as AI at the time, Dan Farmer and Wietse Venema's SATAN program, released in 1995, automated a process for finding vulnerabilities in system configurations that had previously required much more human effort." S. Forrest [10] has proposed a system of inductive reasoning, which belongs to the realm of AI. And one can interpret the work of the group of Vigna et al in UCSB [11], [12], [13], [14], [15]) as proceeding from a somewhat similar philosophy. Arguably Web application firewalls (WAFs) or more generally firewalls using deep packet inspections can be construed as a kind of instantiation of AI in security. Firewalls have been part of the arsenal of cyber-defense for many years. Although more sophisticated techniques are also used, in most cases the filtering is based on port numbers. WAFs cannot rely on port number as most web applications use the same port as the rest of the web traffic. Deep packet inspection is the only option for WAFs to be able to tell a malicious application from a legitimate one. The idea of filtering at the application layer was introduced already in the third generation of firewalls in the 1990's. WAFs can be construed as a special case of application layer firewalls. The modest success of those technologies reflects the need for far more work on AI before they can make significant difference Computer World surveillance.

But those examples show that AI has a natural role to play in Computer World surveillance. They also show that using AI is inherently difficult. Introducing tools whose reaction to attacks is not totally predictable as they involve some context dependent reasoning will force attackers to potentially change completely their strategy. The increasing role of AI is in the logic of the modernization of the web and the internet. Web 2.0, the semantic web, the use of first order logic, the development of OWL are as many evidences of that. These developments will raise the level of complexity of Computer World surveillance and force it to be much more sophisticated. Computer World surveillance may turn out to be one of the best areas of applications of AI.

a) An Illustrative Example: Cross Site Request Forgery (CSRF)

This paragraph is meant to make the following discussion a bit less "high level" or purely abstract, by providing an example around which the rest of the discussion can be organized. The example is Cross Site Request forgery (CSRF). CSRF is not new. In 1988 it was known as "confused deputy". Dubbed a "sleeping giant", it came to prominence (i.e. enter in the OWASP top ten list) in the last few years. In the same way that XSS covers a large spectrum of attacks or vulnerabilities, some of which justifies to be called "cross site scripting", CSRF has grown into an open ended class of attacks. What all these attacks seem to have in common is that they hijack some credentials from a user and use them for the advantage of the attackers. Of this large class of attacks, what follows applies only to a subclass: it is when the attack takes place within the browser of the user. In fact the following considerations would apply to all situations where the attacks take place within the browser of the user, such as the man in the browser. Would an expert monitoring each HTTP request be able in real time to realize that a CSRF attack is unfolding?

Some context dependent analysis is of essence. The decision of whether malicious activity is taking place or not does not need to be the result of only one measurement. The analysis can be protracted and based on a succession of observations. An AI machine using a Bayesian algorithm could potentially do exactly the same. It is known that in a situation where the probability of false positive and negative of individual measurements is not so small, a shrewd Bayesian algorithm, well implemented can dramatically reduce the overall probability of false positive, while maintaining the probability of false negative low [16]. In other words, it is

possible with a shrewd use of multi-observations to maintain the probability of mis-determinations very low [16].

If one takes the example of the CSRF attack described in the classic paper of Felten [17]. A user while in a trusted session with his bank goes to a malicious website and click to download an image. The HTML request is crafted in such a way that through the browser, the query ends going to the bank website and instead of downloading a picture gives instructions (like transferring money or creating a new account) to the bank on behalf of the victim user.

Figure 2.1.1 : General cyber attack

A security tool monitoring the steps would find suspicious to have to find an image at a bank using the trusted session. It would assign a reasonably small probability of false positive for that (which could have been determined statistically before). Parsing the request, it would notice that it carries an executable. This too would raise serious suspicion. Then it could figure in the executable corresponds to what the bank requests for financial transaction. The probability of those three occurrences happening within the same query is small enough to trigger an alert that has a very very small probability of being a false positive. At each step the probability of false positive will be small, but different.

The point is that the tool would have to be able to do those inferences. It should be somewhat intelligent. We now turn to the questions: How far is AI from being able to produce tools like that? And what can be done now to facilitate the development of such tools?

b) A Very Brief Discussion of AI Methods

Although AI can be called a discipline, it is so organized around many different applications that it almost looks like a vast fragmented world. The example of CSRF points toward some form of probabilistic reasoning, as being a more natural approach to deal with that kind of situation than an approach requiring a lot of data for statistical learning for example.

When it comes to probabilistic learning, it seems difficult to avoid contact with Bayesianism. Bayesian reasoning is not without pitfalls, as Judea Pearl intimated in a recent presentation: " I turned Bayesian in 1971, as soon as I began reading Savage's monograph The Foundations of Statistical Inference [18]. The arguments were unassailable: (i) It is plain silly to ignore what we know, (ii) It is natural and useful to cast what we know in the language of probabilities, and (iii) If our subjective probabilities are erroneous, their impact will get washed out in due time, as the number of observations increases.

Thirty years later, I am still a devout Bayesian in the sense of (i), but I now doubt the wisdom of (ii) and I know that, in general, (iii) is false." But (iii) may be false for humans, but not necessarily for machines, because machines do not need to be prejudiced. Algorithms based on Bayesian updating work better with machines than humans.... An additional reason to use a Bayesian approach, is that as we saw in the previous paragraph, through Bayesian updating, it is possible to reduce the probability of false positive and negative. The decision of whether malicious activity is taking place or not does not need to be the result of only one measurement. It can be the result of analyzing a succession of steps.

This vision of a system able to process fast a lot of information, maintaining the rate of false positive and false negative small, despite the fact that the individual components themselves can have a high rate of false positive or false negative is reminiscent of the original idea of von Neumann discussed in his 1956 paper entitled: "Probabilistic logics and the synthesis of reliable organisms from unreliable components" [19].

As can be seen in the CSRF example, this probabilistic reasoning supposes the ability of some autonomous context dependent decision capability, i.e. needs some form of model of the environment. In the words of Judea Pearl [25]: "An intelligent system attempting to build a workable model of its environment cannot rely exclusively on pre-programmed causal knowledge", but must be able to interpret autonomously direct observations. This is where AI differs from more traditional approach to security, which tends to be based on rules. But it is also where the use of AI looks more challenging.

The amount of knowledge virtually present is very much greater than the amount of knowledge explicitly present. The extra knowledge is the result of query-time inference, which can require a lot of computation. And yet, we humans routinely perform this kind of inference quickly and in a way that seems almost effortless." One problem is to access this "virtual knowledge".

We humans also have a remarkable ability that is central to all recognition tasks: we begin with a set of observed features, a set of expectations, and a vast collection of stored descriptions; the problem is to find the stored description that best matches these features and expectations. This is a computationally demanding task that we humans do frequently, quickly, and with no sense of mental effort," but far more challenging for

machines. This is where the AI formidable challenge of identifying algorithms and world representation enters.

The attraction of Knowledge Based systems (KBS) [21] is their ability to make context dependent inferences. KBS tends to be specialized. Knowledge (virtual or real) is acquired or introduced in a variety of ways. But a lot rides on the way knowledge is stored and represented.

Those systems can reason and make inferences. Although they have been developed with different applications in mind, in principle they could be able to make autonomous determination of whether a malicious attack is unfolding.

Computer World surveillance may in fact be an area very appropriate for the application of constructs like the KBS Scone (developed at CMU by Scott Fahlman [22]). Like other KBS, Scone provides support for representing symbolic knowledge about the world: general common-sense knowledge or knowledge about some specific application domain. But Scone is designed to be used as a component in a wide range of software applications. Therefore, a primary emphasis has been put on Scone's expressiveness, ease of use, scalability, and on the efficiency of the most commonly used operations for search and inference. A feature of Scone, which makes it attractive in the context of Computer World surveillance tools is that unlike other KBS (such as Cyc [23], Owl [24], for example, and most Description Logic systems), the emphasis is on the ability to do a lot of simple inference very quickly, not the ability to prove deep theorems or to solve complex logic puzzles. In the system of trade-offs that underlie the development of KBS's, the priority was put on "expressiveness" and "scalability". At this stage, it is far too premature to exclude or recommend any approach.

c) What AI Could Bring to Computer World surveillance

CSRF is meant only as an example. It is only one in an already large and increasing number of web-applications vulnerabilities. Many popular websites are known to have exploitable cross-site scripting (XSS) [2] or cross site request forgery (CSRF, [24], [1]), "ClickJacking" vulnerabilities[4].

Most existing defenses against CSRF are ad hoc. Since CSRF involves in general hijacking a trusted session between a user and a website, a natural approach is to make such hijacking more difficult. One possibility is to not rely excessively on cookies to build trust, but add additional identifiers, at the cost of making trusted sessions more burdensome. A minor consideration in the security community, but the cumulative effect of making every "critical" interaction cumbersome is to project the impression that the logic of the culture of security is just the opposite of the logic of the technological innovations taking place in the internet and the web.

Since users for their security should not have to rely on website designers to anticipate all forms of attacks, tools protecting directly the users are intrinsically more attractive. The user side proposals tend to be based on rules tailored for each known scenario of attack. An example is to treat HTTP POST requests as more dangerous, because they are the one that changes the state of the server and forbid them in some circumstances. In addition to interfere with some useful functionality, this is not sufficient since it is possible (using Java script code injection) to make GET requests accomplish the same thing as POST requests. Disallowing or limiting the use of Java script has been suggested. This makes sense in some specific cases, but that kind of approach is an exacerbated version of security standing on the way of functionalities.

The commercial tool Request Rodeo is a commercial tool, which precludes scenarios on the basis of rules, that could lead to CSRF attacks. Its rules are so strict that it precludes proper interactions with a large number of modern websites. Relaxing the rules on the other hand would reduce the degree of protection, demonstrating if need be that an ideal tool would have to be more intelligent than simply applying rules.

Web application security generates a very new type of challenges compared with the world of worms, buffer overflows, etc.. The two worlds overlap, but they call for very different kinds of security tools. In fact there is no good security tool or even paradigm yet for web application. From the perspective of Computer World surveillance, the world of web applications is very complicated as it seems to offer an infinite numbers of opportunities for abuse.

Some exploitable vulnerabilities are difficult to understand or anticipate as they result from technical details of protocols, implementation of application or are consequences of abusing functionalities which otherwise are very useful or valuable. This is at a time where web applications are proliferating fast and playing an increasingly central role in many critical operations performed in the internet. Some exploitable vulnerabilities are difficult to understand or anticipate as they result from technical details of protocols, implementation of application or are consequences of abusing functionalities which otherwise are very useful or valuable. Most of web application vulnerabilities stems from what makes HTTP, HTML, Java-Script and the like so efficient to support web activity. The controversy around the "same origin policy" illustrates the complication of web application security. All this to show that web application security calls for intelligent tools.

Approaches to defense deliberately relying on AI may not deliver quick results. But they offer the perspective of a future very different and far more attractive than what the present approach based on "tweaking the plumbing" offers. In the same way that the co-evolution of pathogens and defenses from biological

organisms has led to the emergence of the immune system, and AI-based approach to cyberdefense can be seen as the natural next step.

To the legitimate concern that AI-based tools may be very large and suck a lot of CPUs (the immune system has as many cells as the nervous system: it is a huge organ), one can point to the fact that there is not much alternative. One immediate predictable benefit of approaching Computer World surveillance from and AI point of view will be to raise the level of the debate. Instead of being lost in the details of the implementation of old ideas, it will be forward looking.

On the other hand, it has to be recognized also that letting security be the organizing principle of modernization of the internet, will have a stifling effect on innovations, i.e. one of the main reasons of the success of the internet. But if innovation and Computer World surveillance begin to use the common perspective of AI, both at the same time rely increasingly on AI, the logic of the interaction will be dramatically different.

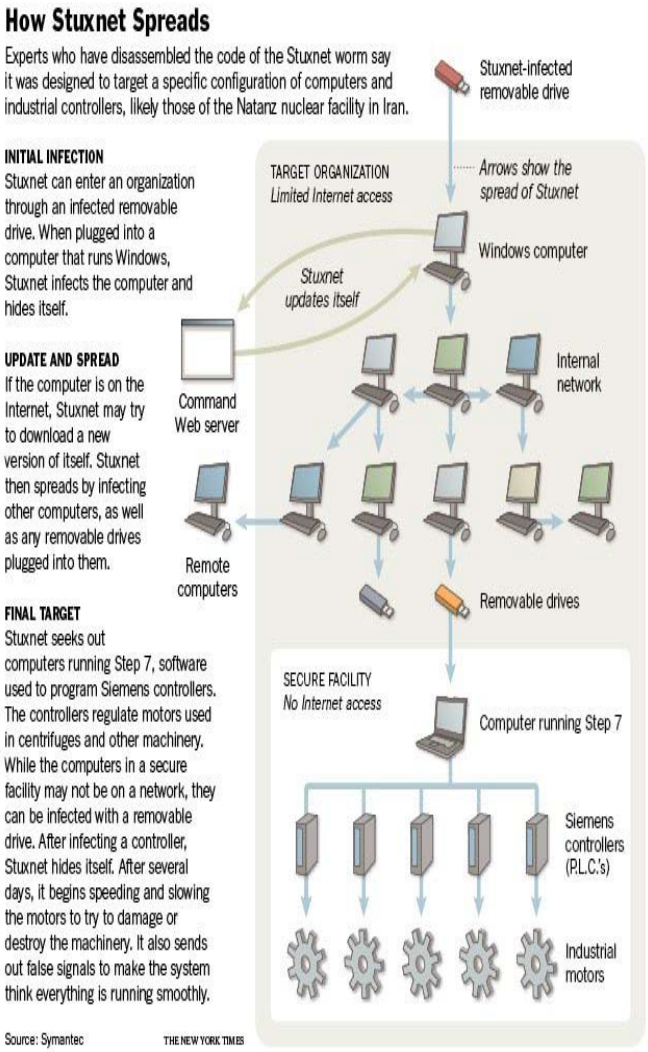


Figure 2.3.1 : How stuxnet Spreads

References Références Referencias

1. Barth, C. Jackson, and J. C. Mitchell. Robust Defenses for Cross-Site Request Forgery. In Proceedings of 15th ACM Conference, CCS, 2008
2. Seth Fogie, Jeremiah Grossman, Robert Hansen, Anton Rager, and Petko D. Petkov. XSS Attacks: Cross Site Scripting Exploits and Defense. Syngress, 2007.
3. N. Jovanovic, E. Kirda, and C. Kruegel. Preventing Cross Site Request Forgery Attacks. Securecommand Workshops, 2006, pages 1–10, Aug. 28 2006- Sept. 1 2006
4. Davide Balzarotti, Marco Cova, Vika Felmetsger, Nenad Jovanovic, Engin Kirda, Christopher Kruegel§, and GiovanniVigna,: Saner: Composing Static and Dynamic Analysis to Validate Sanitization in Web Applications, Proceedings of the IEEE Symposium on Security and Privacy, Oakland, CA, May 2008.
5. Landwehr,Cal, Computer World surveillance and Artificial Intelligence: From Fixing the Plumbing to Smart Water, IEEE, Security and privacy, September/October 2008, p.3
6. Bruce Schneier, On Security, 2008.
7. Douglas Corckford, Ajax Security, 2006.
8. http://www.w3.org/TR/access-control/
9. http://www.owasp.org/index.php/SQL_Injection_Prevention_Cheat_Sheet
10. Kenneth Ingham, Anil Somayaji, John Burge, Stephanie Forrest , Learning DFA representations of HTTP for protecting web applications Journal of Computer Networks. 51:5, pp. 1239-1255 (2007).
11. Darran Mutz, William Robertson, Giovanni Vigna, and Richard Kemmerer, Exploiting Execution Context for the Detection of Anomalous System Calls, Proceedings of the International Symposium on Recent Advances in Intrusion Detection (RAID), Gold Coast Australia, 2007
12. Marco Cova, Davide Balzarotti, Viktoria Felmetsger, and Giovanni Vigna Swaddler : An Approach for the Anomaly-based Detection of State Violations in Web Applications,
13. Robertson, Federico Maggi, Christopher Kruegel Giovanni Vigna, Effective Anomaly Detection with Scarce Training Data, Proceedings of the Network and Distributed System. Security Symposium (NDSS), San Diego, CA, February 010.
14. B. Morel, "Anomaly based intrusion detection systems", chapter in "intrusion detection systems,Intech (2011).
15. William Zeller and Edward W. Felten; Cross-Site Request Forgeries: Exploitation and Prevention, Princeton (2008); http://citp.princeton.edu/csrf/
16. L. J. Savage: The foundations of statistical inferences, 1962.

17. J. von Neumann, "Probabilistic logics and the synthesis of reliable organisms from unreliable components", in C. E. Shannon and J.McCarthy,editors, Annals of Math Studies, numbers 34, pages 43-98. Princeton Univ. Press, 1956
18. Scott Fahlman: "NETL, a System for Representing and Using real World Knowledge", MIT Press, Cambridge, MA, 1979
19. R. Akerkar, P.S. Sajja, Knowledge Based Systems, Jones and Bartlett, 2009.
20. Fahlman, S.E.: The Scone Knowledge Base (homepage), http://www.cs.cmu.edu/~sef/ scone/
21. Blake Shepard et al. (2005). "A Knowledge-Based Approach to Network Security http://www.w3.org/2004/OWL/
22. Judea Pearl. Probabilistic Reasoning in Intelligent systems: Networks of Plausible Inference. Morgan Kaufmann, San Mateo, CA, 1988.

Global Journal of Computer Science and Technology
Neural & Artificial Intelligence
Volume 12 Issue 10 Version 1.0 Year 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

An Empirical Investigation of Using ANN Based N-State Sequential Machine and Chaotic Neural Network in the Field of Cryptography

By Nitin Shukla & Abhinav Tiwari
University of RGPV Bhopal Madhya Pradesh India

Abstract - Cryptography is the exchange of information among the users without leakage of information to others. Many public key cryptography are available which are based on number theory but it has the drawback of requirement of large computational power, complexity and time consumption during generation of key [1]. To overcome these drawbacks, we analyzed neural network is the best way to generate secret key. In this paper we proposed a very new approach in the field of cryptography. We are using two artificial neural networks in the field of cryptography. First One is ANN based n-state sequential machine and Other One is chaotic neural network. For simulation MATLAB software is used. This paper also includes an experimental results and complete demonstration that ANN based n-state sequential machine and chaotic neural network is successfully perform the cryptography.

Keywords : *ANN, n-state Sequential Machine, Chaotic Neural Network, Cryptography.*

GJCST-E Classification: *D.4.6*

Strictly as per the compliance and regulations of:

An Empirical Investigation of Using ANN Based N-State Sequential Machine and Chaotic Neural Network in the Field of Cryptography

Nitin Shukla [α] & Abhinav Tiwari [σ]

Abstract - Cryptography is the exchange of information among the users without leakage of information to others. Many public key cryptography are available which are based on number theory but it has the drawback of requirement of large computational power, complexity and time consumption during generation of key [1]. To overcome these drawbacks, we analyzed neural network is the best way to generate secret key. In this paper we proposed a very new approach in the field of cryptography. We are using two artificial neural networks in the field of cryptography. First One is ANN based n-state sequential machine and Other One is chaotic neural network. For simulation MATLAB software is used. This paper also includes an experimental results and complete demonstration that ANN based n-state sequential machine and chaotic neural network is successfully perform the cryptography.

Keywords : ANN, n-state Sequential Machine, Chaotic Neural Network, Cryptography.

I. Introduction

An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process[12]. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons.

[12]Cryptosystems are commonly used for protecting the integrity, confidentiality, and authenticity of information resources. In addition to meeting standard specifications relating to encryption and decryption, such systems must meet increasingly stringent specifications concerning information security. This is mostly due to the steady demand to protect data and resources from disclosure, to guarantee the authenticity of data, and to protect systems from web based attacks. For these reasons, the development and evaluation of cryptographic algorithms is a challenging task.

This paper is an investigation of using ANN based n-state sequential machine and chaotic neural network in the field of cryptography .the rest of the paper is organized as follows: section 2 discusses background and related work in the field of ANN based cryptography, section 3 proposed method related to n-state sequential machine and chaotic neural network section 4 discusses implementation section 5 discusses experimental report and test result and finally section 6 discusses conclusion.

II. Background and Related Work

Jason L. Wright , Milos Manic Proposed a research paper on Neural Network Approach to Locating Cryptography in Object Code. In this paper, artificial neural networks are used to classify functional blocks from a disassembled program as being either cryptography related or not. The resulting system, referred to as NNLC (Neural Net for Locating Cryptography) is presented and results of applying this system to various libraries are described[2].

John Justin M, Manimurugan S introduced A Survey on Various Encryption Techniques. This paper focuses mainly on the different kinds of encryption techniques that are existing, and framing all the techniques together as a literature survey. Aim an extensive experimental study of implementations of various available encryption techniques. Also focuses on image encryption techniques, information encryption techniques, double encryption and Chaos-based encryption techniques. This study extends to the performance parameters used in encryption processes and analysing on their security issues[3].

Ilker DALKIRAN, Kenan DANIS¸MAN introduced a research paper on Artificial neural network based chaotic generator for cryptology . In this paper, to overcome disadvantages of chaotic systems, the dynamics of Chua's circuit namely x, y and z were modeled using Artificial Neural Network (ANN). ANNs have some distinctive capabilities like learning from experiences, generalizing from a few data and nonlinear relationship between inputs and outputs. The proposed ANN was trained in diffrent structures using different learning algorithms. To train the ANN, 24 different sets including the initial conditions of Chua's circuit were used and each set consisted of about 1800 input-output data. The experimental results showed that a feed-

Author α : Assistant Professor in Computer Science Dept. from S.R.I.T. Jabalpur University of RGPV Bhopal Madhya Pradesh India.
Author σ : M.Tech. Scholar S.R.I.T. Jabalpur University of RGPV Bhopal Madhya Pradesh India.

forward Multi Layer Perceptron (MLP), trained with Bayesian Regulation back propagation algorithm, was found as the suitable network structure. As a case study, a message wasfirst encrypted and then decrypted by the chaotic dynamics obtained from the proposed ANN and a comparison was made between the proposed ANN and the numerical solution of Chua's circuit about encrypted and decrypted messages[5].

Eva Volna ,Martin Kotyrba ,Vaclav Kocian, Michal Janosek developed a Cryptography Based on Neural Network. This paper deals with using neural network in cryptography, e.g. designing such neural network that would be practically used in the area of cryptography. This paper also includes an experimental demonstration[6].

Karam M. Z. Othman , Mohammed H. Al Jammas Introduced Implementation of Neural - Cryptographic System Using Fpga. In this work, a Pseudo Random Number Generator (PRNG) based on artificial Neural Networks (ANN) has been designed. This PRNG has been used to design stream cipher system with high statistical randomness properties of its key sequence using ANN. Software simulation has been build using MATLAB to firstly, ensure passing four well-known statistical tests that guaranteed randomness characteristics. Secondly, such stream cipher system is required to be implemented using FPGA technology, therefore, minimum hardware requirements has to be considered[7].

T. Schmidt , H. Rahnama Developed A Review of Applications of Artificial Neural Networks In Cryptosystems. This paper presents a review of the literature on the use of artificial neutral networks in cryptography. Different neural network based approaches have been categorized based on their applications to different components of cryptosystems such as secret key protocols, visual cryptography, design of random generators, digital watermarking, and steganalysis[8].

Wenwu Yu, Jinde Cao introduced Cryptography based on delayed chaotic neural networks. In this Letter, a novel approach of encryption based on chaotic Hopfield neural networks with time varying delay is proposed. We use the chaotic neural network to generate binary sequences which will be used for masking plaintext. The plaintext is masked by switching of chaotic neural network maps and permutation of generated binary sequences. Simulation results were given to show the feasibility and effectiveness in the proposed scheme of this Letter. As a result, chaotic cryptography becomes more practical in the secure transmission of large multi-media files over public data communication network[9]

III. Proposed Method

A number of studies have already investigated different machine learning methodologies, specifically neural networks and their applications in cryptography, but It is uncommon technique to using Artificial Neural network based n-state sequential machine and Chaotic neural network in the field of cryptography .

a) Sequential Machine

A sequential Machine output depends on state of the machine as well as the input given to the sequential machine. Therefore Michel I .Jordan Network was designed because in which output are treated as input. We are used these type of input as a state.

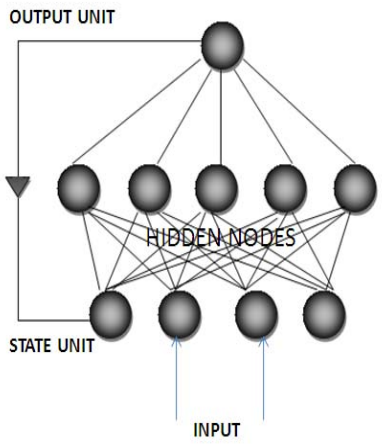


Fig. 3.1 : Michel I .Jordan Neural Network

Multilayer network has been designed with the help of Michel I .Jordan Network Fig 3.1 .In this network has 3 layers an input layer, a hidden layer and an output layer. The size of the input layer depends on the number of inputs and the number of outputs being used to denote the states. The learning algorithm used for this network is back propagation algorithm and the transfer function in the hidden layer is a sigmoid function. For implementation of sequential machine a serial adder and a sequential decoder is used.

b) Cryptography Achieved by Artificial neural network based n-state sequential machine

As a sequential machine can be achieved by using a Michel I. Jordan neural network, therefore data are successfully encrypted and decrypted. In this case the starting state of the n-state sequential machine can act as a key. Data is used to train the neural network as it provides the way the machine moves from one state to another.

c) Cryptography Achieved by a chaotic neural network

Cryptography scheme was done by a chaotic neural network. A network is called chaotic neural network if its weights and biases are determined by chaotic sequence. Specially encryption of digital signal we used chaotic neural network.

IV. Implementation

a) Sequential Machine

A finite state sequential machine was implemented using a Michael I. Jordan network is used. Jordan networks are also known as "simple recurrent networks". Using back propagation algorithm for train Jordan Neural network. The application of the generalized delta rule thus involves two phases:

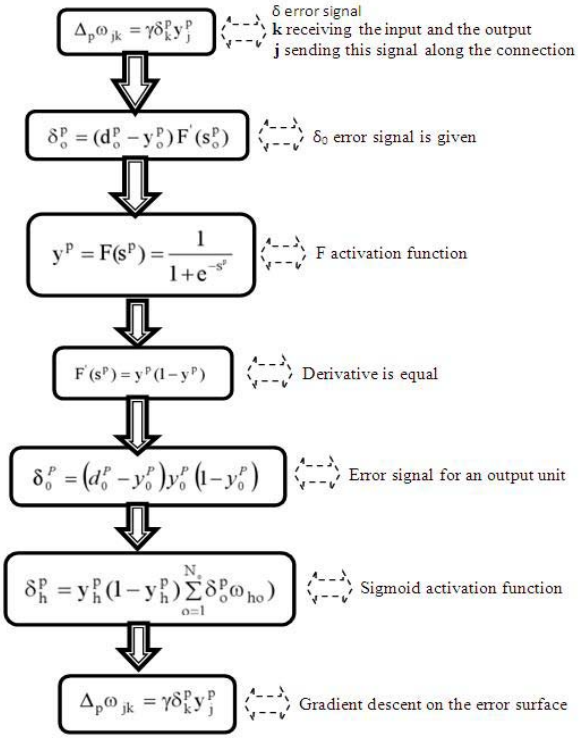


Fig. 4.1 : Weight adjustments with sigmoid activation function

First Phase: The input x is presented and propagated forward through the network to compute the output values yp for each output unit. This output is compared with its desired value do, resulting in an error signal δp for each output unit.

The Second Phase: This involves a backward pass through the network during which the error signal is passed to each unit in the network and appropriate weight changes are calculated.

b) Cryptography achieved by Using ANN based sequential machine

The reason for using sequential machine for implementation is that the output and input can have any type of relationship and the output depends on the starting state. The starting state is used as a key for encryption and decryption. If the starting state is not known, it is not possible to retrieve the data by decryption even if the state table or the working of the sequential state is known. For training of the neural network, any type of sequential machine can be used with the key showing the complexity or the level of security obtained.

c) Cryptography Achieved Through chaotic neural network

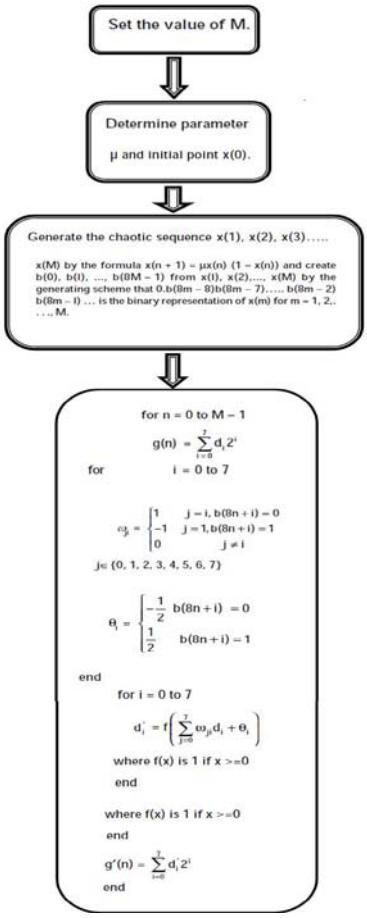


Fig. 4.2 : CNN based Algorithm for encryption

A chaotic neural network in which weights and biases are determined by a chaotic sequence.

g = digital signal of length M and g (n)

0 ≤ M-1 , be the one- byte value of the signal g at position n .The decryption procedure is the same as the above one except that the input signal to the decryption Chaotic neural network should be g'(n) and its output signal should be g"(n).

V. Experiment and Test Result

a) Sequential Machine

A general n-state Sequential Machine was implemented. As an example, the serial adder was implemented using this machine.

Input 1	Input 2	Current state	Output	Next state
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	0	1
1	0	0	1	0
1	0	1	0	1
1	1	0	0	1
1	1	1	1	1

Table 5.1 : State table of the Serial Adder

Table 5.1 demonstrate the state table of the Serial Adder The current state represents any previous carry that might be present whereas the next state represents the output carry. Serial Adder State table data is entered into the program. The following Command window implemented in MATLAB show different stages of the execution.

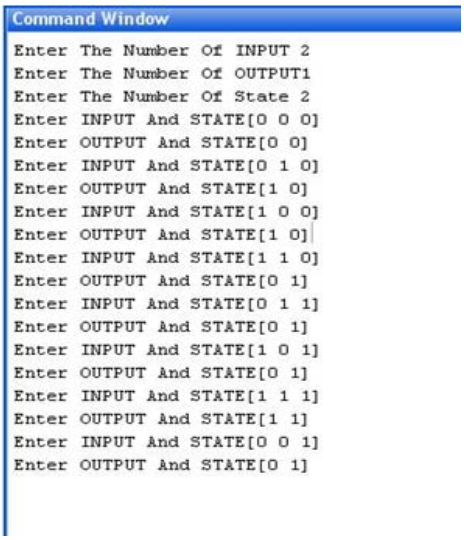


Fig. 5.2 : Enter the training data in the n-State sequential machine

N-state sequential machine program is implemented and enter training data .First it ask user to enter input ,output and state here we enter 2 input, 1 output and 2 states. With the help of back-propagation algorithm, to minimize the error function.

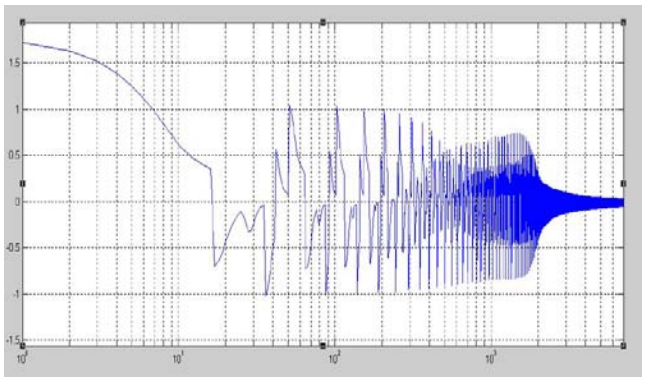


Fig. 5.3 : Shows the plot of the error function against the number of iterations

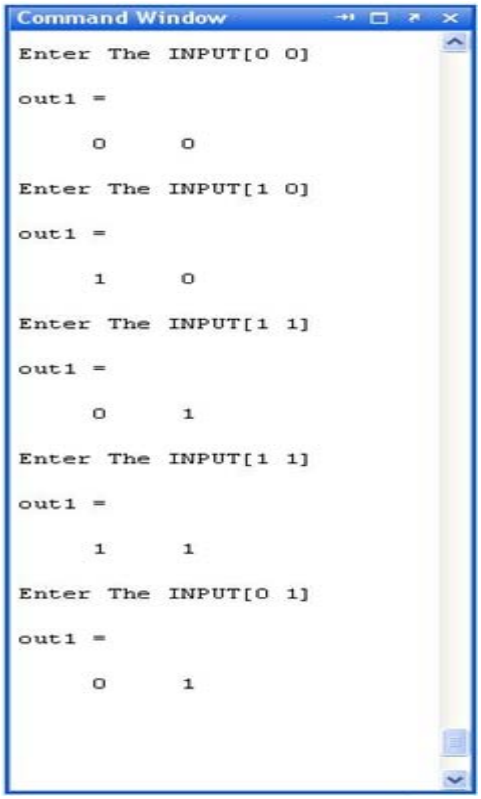


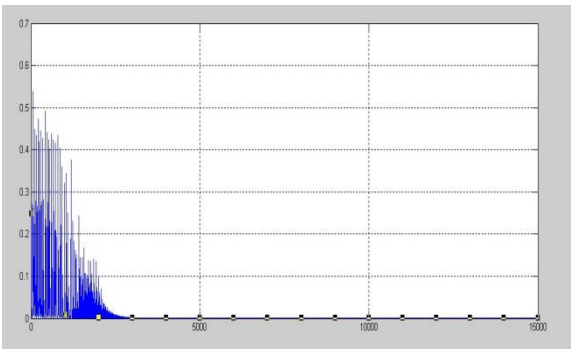
Fig. 5.4 : Final output of sequential machine

Above Command window show Final Output of the sequential machine implemented as a Serial Adder. There is initial state is informed to the user for the input bits to be added. The output is the sum and the carry bit. After that the execution of program has been completed it is automatically jumps to the new carry state. This output is considered as previous carry state.

Following Graph illustrate Mean Square Error when we Enter input 0,1 and 2 .we apply this input in two feed forward Adder One is of Multilayer single output feed forward Adder and other is Multilayer multiple output feed forward Adder.

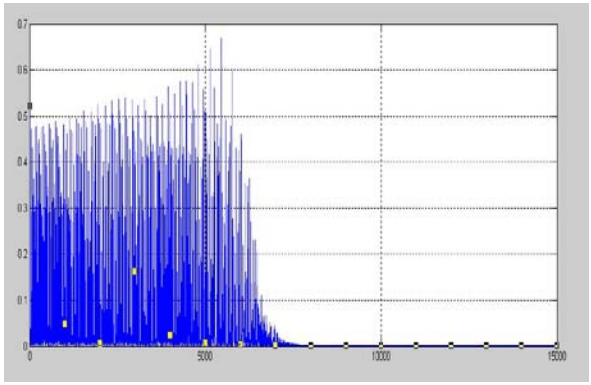
First we apply Input Number 0,1 and 2 on Multilayer single output feed forward and find MSE on Linear scale .

Enter Input 0:



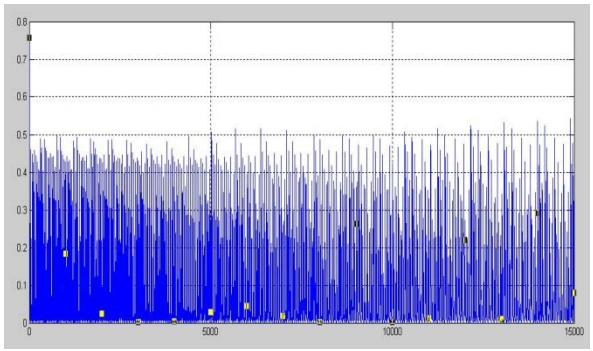
(a) MSE of Multilayer single output feed forward

Enter Input 1:



(b) MSE of Multilayer single output feed forward

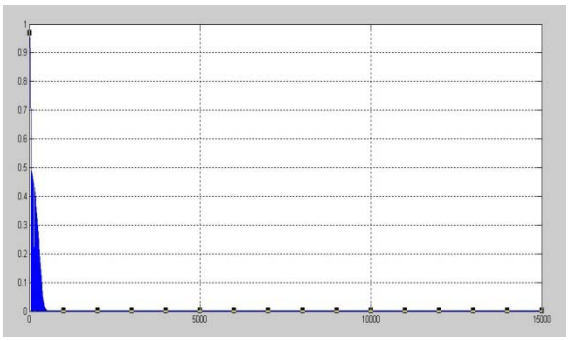
Enter Input 2:



(c) MSE of Multilayer single output feed forward

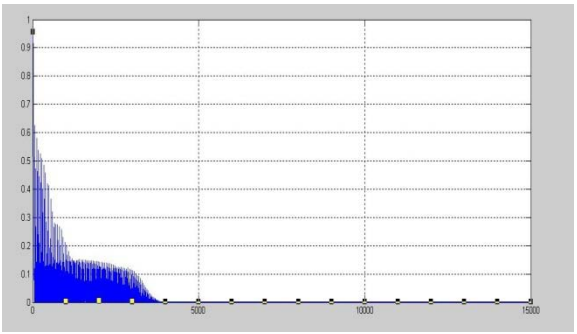
Secondly we apply Input Number 0,1 and 2 on Multilayer multiple output feed forward Adder and find MSE on Linear scale.

Enter Input: 0:



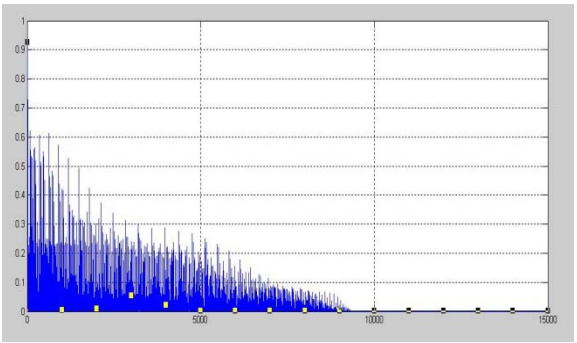
(d) MSE of Multilayer multiple output feed forward

Enter Input : 1:



(e) MSE of Multilayer multiple output feed forward

Enter Input : 2:



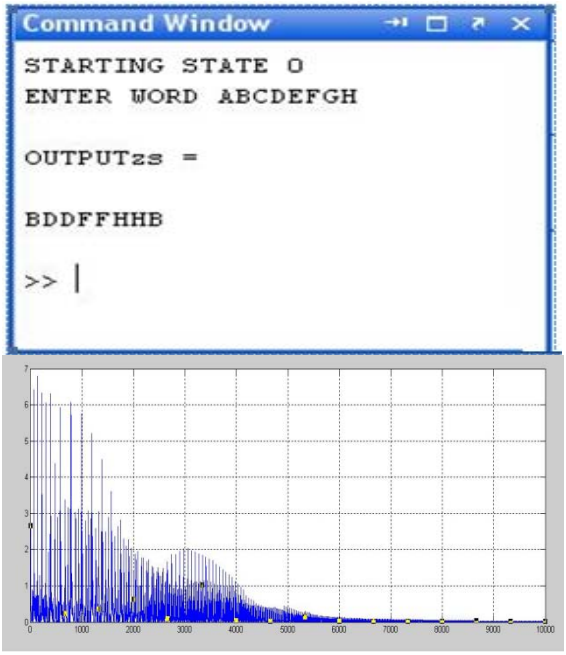
(f) MSE of Multilayer multiple output feed forward

All these graph shows Mean Square Error on linear scale .Comparative study is done between Multilayer single output feed forward Adder and Multilayer multiple output feed forward Adder. Multilayer multiple output feed forward Adder generate smaller number of patterns and thus reduced the training time as well as the number of neurons in compare to and Multilayer single output feed forward Adder.

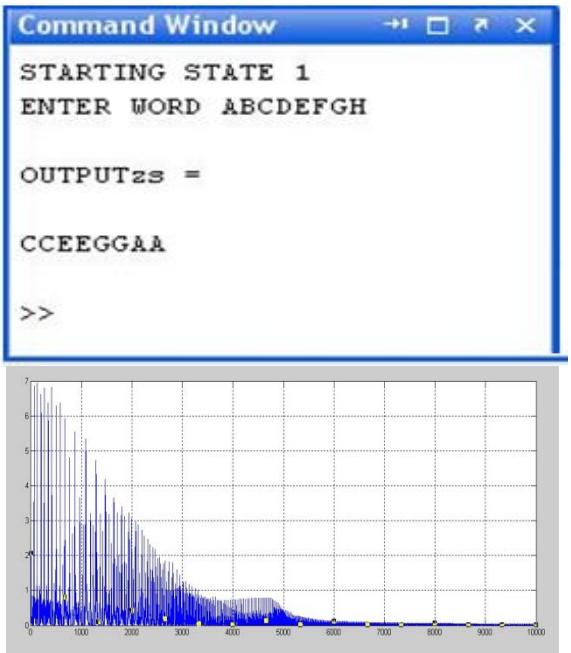
b) Cryptography achieved Through ANN based Sequential Machine

In this section with the help of ANN based n-state sequential machine successfully convert a Letter A to H in Encrypted form. it ask user to enter state, if the starting state is 0 then the input letter is shifted by one

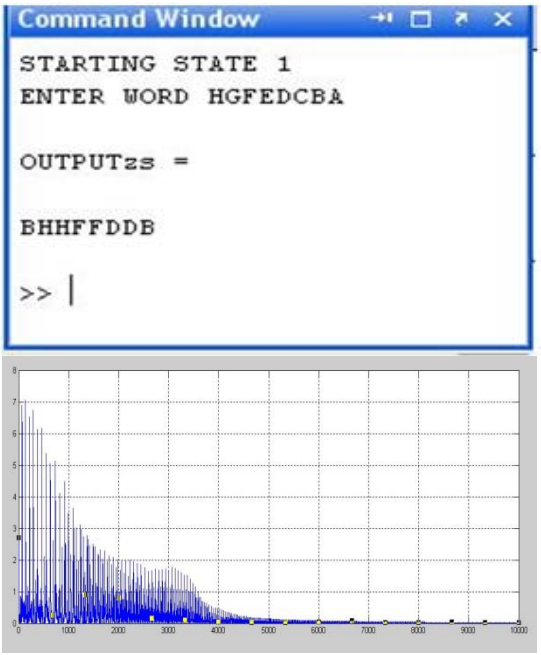
and generated encrypted letter similarly is the starting state is 1 the letter is shifted by 2.The state is automatically move to next state .If the next input is again A the output will be C as the current state now is 1. For H, state 0 will flip the letter to A while state 1 will flip the output to B. This method can be used to encrypt a word containing only the letters A to H.



(a) Output using n-state sequential machine based Encryption on code window and output graph on linear scale starting state 0



(b) Output using n-state sequential machine based Encryption on code window and output graph on linear scale starting state 1



(c) Output using n-state sequential machine based Encryption on code window and output graph on linear scale starting state 1

c) Cryptography Achieved by chaotic neural network

A chaotic network is a neural network whose weights depend on a chaotic sequence. The chaotic sequence highly depends upon the initial conditions and the parameters, x(0) and μ are set. It is very difficult to decrypt an encrypted data correctly by making an exhaustive search without knowing x(0) and $\mu()$.

Table 1 : Same Input Encrypted with Different Initial Conditions (Values of x(0) and μ())

Same Input Encrypted with Different Initial Conditions (Values of x(0) and μ())				
INPUT	ASCII CODE	Output with x(0)= 0.75 μ=3.9	Output with x(0)= 0.85 μ=3.5	output with x(0)= 0.90 μ=3.2
A	97	199	233	204
B	98	195	112	98
C	99	200	239	11
D	100	253	108	31
E	101	220	226	1
F	102	17	115	25
G	103	187	234	56
H	104	101	109	235
I	105	6	236	49
J	106	138	115	226
K	107	107	229	36
L	108	32	110	225
M	109	180	238	42
N	110	119	112	254
O	111	225	224	45
P	112	184	113	225
Q	113	63	243	49
R	114	168	83	227
S	115	103	252	51
T	116	245	116	229
U	117	160	244	53
V	118	83	84	231
W	119	209	248	55
X	120	219	120	233
Y	121	209	248	57
Z	122	231	88	235

Table 2 : Encrypted Data of Table 1 (Column 2) Decrypted Using Same and Different Initial Conditions

Encrypted Data of Table 1 (Column 2) Decrypted Using Same and Different Initial Conditions				
		Output Obtained Using Same Initial Condition	Output Obtained Using Different Initial Condition	
INPUT	ASCII CODE	output with x(0)= 0.75 μ=3.9	output with x(0)= 0.85 μ=3.5	output with x(0)= 0.90 μ=3.2
A	199	97	79	106
B	195	98	209	195
C	200	99	68	160
D	253	100	245	134
E	220	101	91	184
F	17	102	4	110
G	187	103	54	228
H	101	104	96	230
I	6	105	131	94
J	138	106	147	2
K	107	107	229	36
L	32	108	34	173
M	180	109	55	243
N	119	110	105	231
O	225	111	110	163
P	184	112	185	41
Q	63	113	189	127
R	168	114	137	57
S	103	115	232	39
T	245	116	245	100
U	160	117	33	224
V	83	118	113	194
W	209	119	94	145
X	219	120	219	74
Y	209	121	80	145
Z	231	122	197	118

Table 3 : Encrypted Data of Table 1 (Column 3) Decrypted Using Same and Different Initial Conditions

Encrypted Data of Table 1 (Column 3) Decrypted Using Same and Different Initial Conditions				
Output Obtained Using Same Initial Condition		Output Obtained Using Different Initial Condition		
INPUT	ASCII CODE	output with x(0)= 0.75 μ=3.9	output with x(0)= 0.85 μ=3.5	output with x(0)= 0.90 μ=3.2
A	233	79	97	68
B	112	209	98	112
C	239	68	99	135
D	108	245	100	23
E	226	91	101	134
F	115	4	102	12
G	234	54	103	181
H	109	96	104	238
I	236	131	105	180
J	115	147	106	251
K	229	229	107	170
L	110	34	108	227
M	238	55	109	169
N	112	105	110	224
O	224	110	111	162
P	113	185	112	224
Q	243	189	113	179
R	83	137	114	194
S	252	232	115	188
T	116	245	116	229
U	244	33	117	180
V	84	113	118	197
W	248	94	119	184
X	120	219	120	233
Y	248	80	121	184
Z	88	197	122	201

Table 4 : Encrypted Data of Table 1 (Column4) Decrypted Using Same and Different Initial Conditions

Encrypted Data of Table 1 (Column 4) Decrypted Using Same and Different Initial Conditions				
Output Obtained Using Same Initial Condition		Output Obtained Using Different Initial Condition		
INPUT	ASCII CODE	output with x(0)= 0.75 μ=3.9	output with x(0)= 0.85 μ=3.5	output with x(0)= 0.90 μ=3.2
A	204	106	68	97
B	98	195	112	98
C	11	160	135	99
D	31	134	23	100
E	1	184	134	101
F	25	110	12	102
G	56	228	181	103
H	235	230	238	104
I	49	94	180	105
J	226	2	251	106
K	36	36	170	107
L	225	173	227	108
M	42	243	169	109
N	254	231	224	110
O	45	163	162	111
P	225	41	224	112
Q	49	127	179	113
R	227	57	194	114
S	51	39	188	115
T	229	100	229	116
U	53	224	180	117
V	231	194	197	118
W	55	145	184	119
X	233	74	233	120
Y	57	145	184	121
Z	235	118	201	122

It is clear from table 2, 3 and 4 that we can decrypt an encrypted data correctly by knowing the exact values of x (0) and μ otherwise we get the wrong data as shown in table 2,3 and 4.

VI. Conclusion

Our experiments lead to the following conclusions.

1. Form the experiment section 4 (A) it clear that Sequential Machine was successfully trained with the help back propagation algorithm of ANN. With the help of back-propagation algorithm, to minimize the error function. We also compare same inputs passes between two feed forward adders. Our experiment and test result was also showing Mean square Error between them. Multilayer multiple output feed forward adder show better result as compare to Multilayer multiple output feed forward Adder. Multilayer multiple output feed forward Adder generate smaller number of patterns and thus reduced the training time as well as the number of

neurons in compare to and Multilayer single output feed forward Adder.

2. In the experiment section 4 (B) it is clears that ANN based n- state sequential machine successfully con-vert a Letter A to H in Encrypted form.
3. In the experiment section 4 (c) it is clears from table 2, 3 and 4 that we can decrypt an encrypt data correctly by knowing the exact values of x (0) and μ otherwise we get the wrong data . ASCII CODE decimal value of alphabet A to Z securely encrypted and decrypted using chaotic neural network. Test result and related graph clearly identified parameter on which training time reduces. Artificial neural network successfully built and trained sequential machine and chaotic neural network for performing cryptography.

References Références Referencias

1. Shweta B. Suryawanshi and Devesh D. Nawgaje- a triple-key chaotic neural network for cryptography in image processing International Journal of Engineering Sciences & Emerging Technologies, April 2012. ISSN: 2231 – 6604 Volume 2, Issue 1, pp: 46-50 ©IJESET
2. E.C. Laskari , G.C. Meletiou, D.K. Tasoulis , M.N. Vrahatis , Studying the performance of artificial neural networks on problems related to cryptography , Nonlinear Analysis: Real World Applications 7 (2006) 937 – 942
3. Jason L. Wright , Milos Manic - Neural Network Approach to Locating Cryptography in Object Code. Emerging Technologies and Factory Automation INL Laboratory.
4. John Justin M, Manimurugan S - A Survey on Various Encryption Techniques , International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2, Issue-1, March 2012
5. Harpreet Kaur , Tripatjot Singh Panag CRYPTOGRAPHY USING CHAOTIC NEURAL NETWORK International Journal of Information Technology and Knowledge Management July-December 2011, Volume 4, No. 2, pp. 417-422
6. Ilker DALKIRAN, Kenan DANIS¸MAN - Artificial neural network based chaotic generator for cryptology ,Turk J Elec Eng & Comp Sci, Vol.18, No.2, 2010, © T¨UB˙ITAK
7. Eva Volna ,Martin Kotyrba ,Vaclav Kocian,Michal Janosek - CRYPTOGRAPHY BASED ON NEURAL NETWORK , Department of Informatics and Computers University of Ostrava Dvorakova 7, Ostrava, 702 00, Czech Republic.
8. KARAM M. Z. OTHMAN , MOHAMMED H. AL JAMMAS - IMPLEMENTATION OF NEURAL - CRYPTOGRAPHIC SYSTEM USING FPGA . journal of Engineering Science and Technology Vol. 6, No. 4 (2011) 411 – 428 © School of Engineering, Taylor's University
9. Wenwu Yu, Jinde Cao - Cryptography based on delayed chaotic neural networks Department of Mathematics, Southeast University, Nanjing 210096, China Received 1 February 2006; received in revised form 10 March 2006; accepted 28 March 2006 Available online 17 April 2006Communicated by A.R. Bishop
10. Shweta B. Suryawanshi and Devesh D. Nawgaje- a triple-key chaotic neural network for cryptography in image processing International Journal of Engineering Sciences & Emerging Technologies, April 2012. ISSN: 2231 – 6604 Volume 2, Issue 1, pp: 46-50 ©IJESET
11. Jay Kumar Ankit Sinha ,Guncha Goswami, Manisha Kumari ,Ratan Singh - Modeling and Simulation of Backpropogation Algorithm Using VHDL International Journal of Computer Applications in Engineering Sciences [VOL I, ISSUE II, JUNE 2011] [ISSN: 2231-4946]
12. T. SCHMIDT, dept. of computer science, ryerson university, canada - a review of applications of artificial neural networks in cryptosystems
13. Srividya, G.; Nandakumar, P, 'A Triple-Key chaotic image encryption method', Communications and Signal Processing (ICCSP), 2011 International Conference on Feb. 2011 ,266 – 270
14. T.Godhavari, 'Cryptography using neural network', IEEE Indicon 2005 Conference, Chennai, India, 11-13 Dec. 2005,258-261.
15. Miles E. Smid and Dennis K. Branstad. 'The Data Encryption Standard: Past and Future', proceedings of the ieee, vol. 76, no. 5, may 1988,550-559.
16. Shweta B Suryawanshi and Devesh D Nawgaje., 'Chaotic Neural Network for Cryptography in Image Processing'. IJCA Proceedings on 2nd National Conference on Information and Communication Technology NCICT(3):, November 2011. Published by Foundation of Computer Science, New York, USA.
17. "Design and Realization of A New Chaotic Neural Encryption/Decryption Network" by Scott Su, Alvin Lin, and Jui-Cheng Yen.
18. "Capacity of Several Neural Networks With Respect to Digital Adder and Multiplier" by Daniel C. Biederman and Esther Ososanya.
19. "Artificial Intelligence A Modern Approach" by Stuart J. Russell and Peter Norvig.
20. Digital design, fourth edition by M. Morris Mano and Michael D. Ciletti
21. "Machine Learning, Neural and Statistical Classification" by D. Michie, D.J. Spiegelhalter, C.C. Taylor February 17, 1994.
22. Wasserman, Philip D. Neural Computing, Theory and Practice. Van Nordstrand Reinhold, New York. 1989.

23. M. E. Smid and D. K. Branstad, "The Data Encryption Standard: Past and Future," Proceedings of The IEEE, vol. 76, no. 5, pp. 550-559, 1988.
24. C. Boyd, "Modem Data Encryption," Electronics & Communication Journal, pp. 271-278, Oct. 1993. 131 N. Bourbakis and C. Alexopoulos, "Picture Data Encryption Using SC4N Pattern," Pattern Recognition, vol. 25, no. 6, pp. 567-581, 1992.
25. J. C. Yen and J. I. GUO, "A New Image Encryption Algorithm and Its VLSI Architecture," 1999 IEEE Workshop on Signal Procs. Systems, Grand Hotel, Taipei, Taiwan, Oct. 18-22, pp. 430-437, 1999.

Global Journal of Computer Science and Technology
Neural & Artificial Intelligence
Volume 12 Issue 10 Version 1.0 Year 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Unsolved Tricky Issues on COTS Selection and Evaluation

By Zahid Javed, Ahsan Raza Sattar & Muhammad Shakeel Faridi
University of Agriculture, Faisalabad, Pakistan

Abstract - Component Based Software Engineering (CBSE) approach is based on the idea to develop software systems by selecting appropriate components and then to assemble them with a well-defined software architecture. (CBSE) offers developers the twin benefits of reduced software life cycles, shorter development times , saving cost and less effort as compare to build own component. However the success of the component based paradigm depends on the quality of the commercial off-the-shelf (COTS) components purchased and integrated into the existing software systems. It is need of the time to present a quality model that can be used by software programmer to evaluate the quality of software components before integrating them into legacy systems. The evaluation and selection of the COTS components are the most critical process. These evaluation and selection method cannot be resolved by the IT professionals itself. In this study the author tried to compare the twenty three available systematic methods for best evaluation and selection of COTS components.

Keywords : Component Based Software Engineering, Commercial off-the-Shelf, Software Architecture.

GJCST-E Classification : D.2.11

Strictly as per the compliance and regulations of:

Unsolved Tricky Issues on COTS Selection and Evaluation

Zahid Javed [α], Ahsan Raza Sattar [σ] & Muhammad Shakeel Faridi [ρ]

Abstract - Component Based Software Engineering (CBSE) approach is based on the idea to develop software systems by selecting appropriate components and then to assemble them with a well-defined software architecture. (CBSE) offers developers the twin benefits of reduced software life cycles, shorter development times , saving cost and less effort as compare to build own component. However the success of the component based paradigm depends on the quality of the commercial off-the-shelf (COTS) components purchased and integrated into the existing software systems. It is need of the time to present a quality model that can be used by software programmer to evaluate the quality of software components before integrating them into legacy systems. The evaluation and selection of the COTS components are the most critical process. These evaluation and selection method cannot be resolved by the IT professionals itself. In this study the author tried to compare the twenty three available systematic methods for best evaluation and selection of COTS components.

Keywords : Component Based Software Engineering, Commercial off-the-Shelf, Software Architecture.

I. Introduction

Now a day, COTS (commercial off-the-shelf) are widely used in current software development. They are pieces or templates of software that can reused for future concern [1, 2]. COTS can be word processors and email packages etc. [3]. Its selection plays a crucial role in development of final/end product [4]. Selection of COTS means check whether a component is fit or not for a required product [5]. Many challenges and efforts are made for COTS selection process during last decades but no effective solution can be produced or developed which we can say Silver built for it [6]. Different solutions are introduced in different conditions for COTS selection and evaluation.

The objective of this review is as following;

- To evaluate the best technique to find out COTS Selection components.
- To identify currently used decision making practices of COTS Evaluation & Selection.
- Impact of COTS component on developers.
- Problematic COTS integrated with exiting system.

Author α : Dept of Computer Science, University of Agriculture, Faisalabad, Pakistan. E-mail : zahidjaved.uaf@gmail.com
Author σ : Dept of Computer Science, University of Agriculture, Faisalabad, Pakistan. E-mail : uaf_raza@hotmail.com
Author ρ : Dept of Computer Science, University of Agriculture, Faisalabad, Pakistan. E-mail : shakeelfaridi@gmail.com

II. The COTS Selection Process

There is no certified method available for COTS selection [6], some repeated methods are defined. Figure 1 below is showing the General COTS selection process;

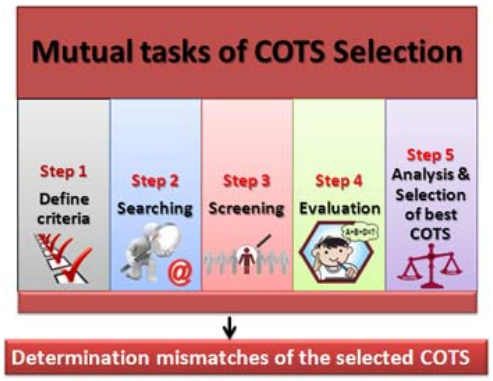


Figure 1 : The General COTS selection (GCS) process

Step1: Evaluation criteria should be defined according to stakeholder's requirements.
Step2: COTS product Selection.
Step3: Filtered resultant components based on requirements.
Step4: Short listed COTS then evaluated.
Step5: Analyze evaluated COTS for best fitness. Normally analytic hierarchy process (AHP), used for selection process [7].

After the final step 5 selections of COTS is done to avoid mismatch problem.

III. COTS Selection Approaches

COTS selection and its strategies are compared here. This section shows how different approaches can contribute during selection of COTS.

a) The Evolution of COTS Selection Practices

First proposed by OTSO (Off-The-Shelf Option) approach for COTS selection in 1995 [16]. OTSO was a milestone towards COTS selection where basic structure was defined. Structure was very like to the GCS process described in Figure 1.

In 1996, Kontio published several followup and papers to elaborate OTSO (e.g.) progressive filtering; evaluation criteria includes functionality, non-functional properties, strategic considerations and architecture

compatibility; AHP suggested for comparison.[10,33]

Many approaches were proposed till 1997. These were

i. The IusWare (IUStitiasoft WAR) is addressed selection process and quality evaluation requirements [17].
ii. The PRISM (Portable, Reusable, Integrated, Software Modules) an architecture proposed that can be effective during COTS evaluation [18].
iii. The CISD (COTS-based Integrated Systems Development) model proposed when multiple homogeneous COTS products was required [19].

In 1998 another approach was introduced named PORE. PORE got importance by proposed requirement engineering process for COTS development. It stated that when COTS are evaluated requirement should gather and analyzed [4].

In 1999, several approaches were proposed:

i. The CEP (Comparative Evaluation Process) approach introduced the use of the socalled confidence factor (CF). The more reliable the source of data, the higher a CF value that source gets. Any estimate we make should be adjusted based on the CF value of the source based on which these estimations are made.
ii. The STACE (Social-Technical Approach to COTSE valuation) approach emphasized the importance of non-technical issues, e.g. social, human, and organizational characteristics, during the evaluation process [20].
iii. The CRE (COTS-based RE) approach emphasized the importance of non-functional requirements (NFR) as decisive criteria when comparing COTS alternatives [21].

In 2000, the COTS acquisition process (CAP) which was an evaluation process. This evaluation process (including the evaluation criteria themselves) should be modified based on the effort available. Ochs approach fits the process using expert knowledge [22].

In 2001 a COTS-Aware Requirements Engineering (CARE) approach was introduced [23-26]. CARE used requirements set according to different agents view.

Another set of approaches were introduced in 2002:

i. The PECA (Plan, Establish, Collect, and Analyze) approach from SEI described COTS selection and where to fit that process [27,5].
ii. The BAREMO approach showed how decision can made using AHP method [7].
iii. The storyboard approach advice to apply use case and other visual methods while requirement gathered from customer, and thus get more fitting COTS products [28].
iv. The mutual selection approach aims to select several COTS that evaluated, initial on the narrow level to evaluate each COTS in separation from the others, and then on the overall level to select the finest combination of COTS [29].
v. i-MATE spotlight on middleware selection and the key role is the narrative of reusable requirements for that area [35].

Two more approaches were proposed:

i. The WinWin spiral model risk management approach that can identify, analyze and mitigate risk [3].
ii. Fuzzy logic approach to produce optimal and quantified solutions [30].

In 2004, the Des COTS system, system integrates some tools to classify evaluation criterion using quality models such as ISO/IEC9126 [31,13].

In 2005, the

i. MiHOS (Mismatch-Handling aware COTS Selection) approach was built up [32]. MiHOS relies on the GCS process which handles mismatch issues between requirements and COTS. MiHOS used method like linear programming to make out optimal way out.
ii. Agile COTS Selection method, Some agility concepts were discussed for COTS selection [38].

In 2011

i. CSSP (COTS Software Selection Process) used by organization and software houses [24].
ii. UnHOS COTS-faced uncertainty issues. Uncertain COTS information can be completeness, accuracy, and consistency. Leaving these uncertainty issues will effect COTS quality and stakeholders satisfaction .Figure 2 shows current COTS selection methods tackle the uncertainty dispute and whether these approaches are supported by a tool or not. [39]

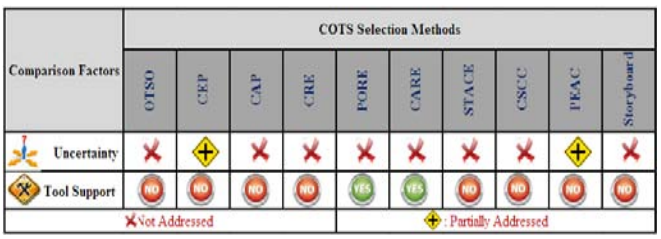


Figure 2 : Comparison of COTS Selection Methods with respect to uncertainty

According to Figure 1, current COTS selections not at all tackle uncertainty. None of these selection methods judge uncertainty in an inclusive style. Only two methods (i.e. PORE and CARE) are hold by a prototype tool. The unavailability of a software tool supporting other COTS selection methods negatively pressures their usability.

There are 23 approaches compare in detail in Figure 3. These approaches in terms of the following criteria:

1) GCS: General COTS Selection method.
2) EVAL: Evaluation strategy used.
3) SNG: Suitability for single COTS selection.
4) MLT: Suitability for multiple COTS selection.
5) MISM: Ability to address COTS mismatches in a systematic way during/after the selection process.
6) TAILOR: Tailor ability of the process based on experts' knowledge. Satisfying this criterion does not necessarily imply the existence of any systematic tailoring techniques.
7) TS: Availability of tool support to facilitate the application of the approach
8) UnHOS: (Uncertainty Handling in COTS Selection) for evaluating COTS candidates while explicitly representing uncertainty. Completeness, accuracy, and consistency

Approach			Comparison Factors							
Sr No.	Name	Year	GCS	EVAL	SNG	MLT	MISM	TAILOR	TS	Uncertainty
1	OTSO	1995	✓	PF	✓	✗	✗	✗	✗	✗
2	IusWare	1997	✓	Any	✓	✗	✗	✗	✗	✗
3	PRISM	1997	✓	PF	✓	✗	✗	✗	✗	✗
4	CISD	1997	✓	PF/PZ	✓	⊕	✗	✗	✗	✗
5	PORE	1998	✓	PF	✓	✗	⊕	✗	✓	✗
6	CEP	1999	✓	PF	✓	✗	✗	✗	✗	⊕
7	STACE	1999	✓	KS/PF	✓	✗	⊕	✗	✗	✗
8	CRE	1999	✓	PF	✓	✗	⊕	✗	✗	✗
9	CAP	2000	✓	PF	✓	✗	✗	✓	✗	✗
10	CARE	2001	✓	PF	✓	✗	⊕	✗	✓	✗
11	RCPEP	2001	✓	PF	✓	✗	✗	✗	✓	✗
12	PECA	2002	✓	Any	✓	✗	⊕	✓	✗	⊕
13	BAREMO	2002	Step5	⊕	✓	✗	✗	✗	✓	✗
14	Storyboard	2002	✗	KS/PF	⊕	✓	⊕	✗	✗	✗
15	CS	2002	✓	⊕	⊕	✓	⊕	✗	✗	✗
16	i-MATE	2002	✓	PF	✓	✗	✗	✗	✗	✗
17	WinWin	2003	✓	PF	✓	⊕	⊕	✗	✗	✗
18	Erol's	2003	✓	⊕	✓	✗	✗	✗	✗	✗
19	DesCOTS	2004	✓	PF	✓	✗	✗	✗	✓	✗
20	Agile	2005	✓	PF	✓	✗	✗	✗	✗	✗
21	MiHOS	2005	✓	KS/PF	✓	✗	✓	✗	✓	✗
22	CSSP	2011	✓	PF	✓	✗	✗	✗	✗	✗
23	UnHOS	2011	✓	PF	✓	✗	✗	✗	✓	✓
PF: Progress Filtering			✓	Full Satisfies the criterion						
KS: Keystone			✗	Does not satisfy the Criterion						
PZ: Puzzle assembly			⊕	Partially, informally, or implicitly satisfies the criterion						

Figure 3 : Comparing COTS selection approaches

IV. COTS Evaluation

It is core for COTS selection because fitness of COTS product based on it. Necessary information is provided in COTS evaluation so select the best fit as per requirement [8, 9].COTS products are evaluated on the base of stakeholders requirements. Suggested hieratically COTS evaluation method in which goals refined according to application requirements and architecture [10]. The practical literature having 6 steps for COTS evaluation according to quality Model [11,12].

Based on the ISO/IEC 9126 quality model [13] the activities defined for COTS evaluation. Three strategies that can applied to evaluate COTS [14, 15]

1. Progressive filtering, Start with large number of COTS and then used iteratively evaluation method, LOW fitted eliminated during each loop. 1to 4 steps used in this strategy in the GCS process repeatedly as desired COTS product is available for system integration.
2. Puzzle assembly, suppose that COTS is like puzzle pieces. This means a COTS product feels bets fit when not integrated but fails to integrate. This shows COTS should be considered in isolation as well as in integration scenario.
3. Keystone identification, starts with requirements identification (e.g. vendor location or type of technology), and searching COTS that fulfills requirements. This enables elimination of not required COTS. One or more strategies are enabled in some projects [15]. Some developer might use 'keystone identification' first and then later 'progressive filtering.

V. Evaluating Existing Approaches

Some issues that are not properly discussed latterly are focused here. Above comparison shows there are varieties for COTS selection. This is open research option:

Problem 1: Best evaluation technique for COTS selection. COTS Selection and evaluation can have following issues.

1. COTS Integration.
2. Mismatch of non-functional requirement for COTS.
3. Indecision Handling (completeness, accuracy, and consistency)
4. Rotating concepts of Stakeholders.
5. Multi Criteria decision-making (MCDM), need for COTS Component.

Different COTS evaluation methods are proposed for different domains. There is no such method available that give solution for above problems. COTS selection is purely base on stakeholders requirements.

Problem 2: Identification of currently COTS selection and evaluation decision making methods. Decision making approaches like Weighted Averages, Fuzzy logic, Bayesian Belief Networks (BBN), Analytic Hierarchy Process (AHP) and linear programming are available. These approaches can apply on Quality selection and evaluation process. IFCOTS software finds then use BBN or AHP, IF COTS component then we used Fuzzy logic or linear programming is used.

Problem 3: COTS components impact on developers.

1. Usually source code of COTS product is not given to developer. This means decision of buy and build

is not easy. Major disadvantage of COTS, it is "black boxes" means not easy to test.

2. COTS component may be mismatched to current system while integrated. Developer should always keep in mind requirements of stakeholder to avoid this situation. Normally COTS product has it specific attributes so it will make misfit while integration to existing system. Two types of mismatches are: Architectural mismatches, and COTS mismatch [36, 37].

Problem 4: Interpretability is another issues occurs when COTS mismatched. This means COTS components mismatches due to lack of adapting quality model while selecting COTS. Using quality model COTS having same standard can match and mismatching can be reduced.

VI. Conclusion

In this paper, we explored the evolution of COTS selection practices, and compared the 23 of the most significant COTS selection approaches. In spite of the great variety of these approaches, there still many open issues related to the COTS selection process that need further research. The objective of this study was to evaluate the best technique to find out COTS Selection components, to identify currently used decision making practices of COTS Evaluation and Selection, impact of COTS component on developers and problematic COTS integrated with exiting system. In future, the author would like to present these models quantitative by using self-completion questionnaire method. This questionnaire method will help out the IT professionals to determine which COTS approach is the best for evaluation and selection of desired components.

References Références Referencias

1. D. G. Firesmith, "Achieving Quality Requirements with Reused Software Components: Challenges to Successful Reuse," inMPEC'05, St. Louis, Missouri, 2005.
2. M. R. Vigder, W. M. Gentleman, & J. Dean, "COTS Software Integration: State of the art," NRC, 39198, Jan 1996.
3. B. Boehm, D. Port, and Y.Yang, "Win Win Spiral Approach to Developing COTSBased Applications, "EDSER-5, Oregon, 2003.
4. N. A. Maiden & C.Ncube, "Acquiring COTS Software Selection Requirements, "IEEE Software, vol.15(2), 1998.pp. 46-56.
5. SEI: Software Engineering Institute in Carnegie Mellon Univ, athttp://www.sei.cmu.edu.
6. G. Ruhe, "Intelligent Support for Selection of COTS Products", LNCS, Springer, vol. 2593 2003. pp. 34-45
7. T. L. Saaty, The analytic hierarchy process. New York: McGraw-Hill, 1990.
8. D. Carney, "Evaluation of COTS Products: Some Thoughts on the Process," SEI Interactive, CMU university, Sept 1998,
9. D. Carney, "COTS Evaluation in the Real World," SEI Interactive, Carnegie Mellon University, Dec 1998,
10. J. Kontio, G. Caldiera, and V. R. Basili, "Defining factors, goals and criteria for reusable component evaluation," in CASCON'96Toronto, Ontario, Canada: IBM Press, 1996.
11. X.Franch &J.P. Carvallo, "Using Quality Models in Software Package Selection, "Software, vol.20(1), Jan/Feb 2003. pp. 34-41
12. J. P. Carvallo, X.Franch, G. Grau, &C. Quer, "COSTUME: a method for building quality models for composite COTS-based software systems," in QSIC'04, Germany, 2004, pp. 214 – 221
13. ISO/IEC1926-1, Software Engineering - Product Quality Model-Part1: Quality Model. Geneve, Switzerland: ISO, 2001.
14. D. Kunda and L. Brooks, "Identifying and Classifying Processes(traditional and soft factors) that Support COTS Component Selection: A Case Study," ECIS'00, Austria, 2000.
15. P. A. Oberndorf, L. Browns word, E. Morris, and C. Sledge," Workshop on COTS Based Systems," SEI Institute, CMU, Special Report CMU/SEI-97-SR-019, Nov. 1997,
16. J. Kontio, "OTSO: A Systematic Process for Reusable Software Component Selection," University of Maryland, Maryland, CSTR- 3478, December 1995,
17. M. Morisio and A. Tsoukis, "IusWare: a methodology for the evaluation and selection of software products," IEE Software Engineering vol. 144 (3), June 1997.
18. R. W. Lichota, R. L. Vesprini, and B. Swanson, "PRISM: Product Examination Process for Component BasedDevelopment," SAST '97, 1997, pp. 61-69.
19. V.Tran&D.Liu,"A Procurement-centric Model for Engineering Component-based Software Systems," SAST '97, 1997,pp.7079.
20. D. Kunda & L.Brooks, "Applying social technical approach for COTS election,"UKAIS'99, Univ of York, McGraw Hill, 1999.
21. L. Chung, B. A. Nixon, E.Yu, and J. Mylopoulos, Non-Functional Requirements in Software Engineering: Kluwer Academic, 99.
22. M.Ochs, D.Pfahl, G. Chrobok- Diening, and B. Nothhelfer-Kolb, "A COTS Acquisition Process: Definition and Application Experience," ESCOMSCOPE' 00, Munich, Germany, 2000.
23. L. Chung and K. Cooper, "A COTS-Aware Requirements Engineering (CARE) Process: Defining System Level Agents, Goals, and Requirements," Department of Computer Science, The University of Texas, Dallas, TR UTDCS-23-01, Dec 2001.

24. L. Chung and K. Cooper, "Knowledge-based COTS-aware requirements engineering approach, "SEKE'02, Ischia, Italy, 2002.
25. L. Chung and K. Cooper, "Defining Goals in a COTS-Aware Requirements Engineering Approach," Systems Engineering journal, vol. 7 (1), 2004, pp. 61-83.
26. L. Chung and K. Cooper, "Matching, Ranking, and Selecting COTS Components: A COTS-Aware Requirements Engineering Approach," in MPEC'04, Scotland, UK, 2004.
27. S. Comella-Dorda, J. C. Dean, E. Morris, and P. Oberndorf, "A Process for COTS Software Product Evaluation," in ICCBSS'02,Orlando, Florida, 2002, pp. 86 - 96.
28. S.Gregor, J.Hutson, & C.Oresky, "Storyboard Process to Assist in Requirements erification and Adaptation to Capabilities Inherent in COTS," in ICCBSS'02, Florida, 2002, pp.132-141.
29. X. Burgues, C. Estay, X. Franch, J. A. Pastor, and C. Quer,"Combined Selection of COTS Components," in ICCBSS'02,Orlando, Florida, 2002, pp. 54-64.
30. I. Erol and W. G. Ferrell-Jr., "A methodology for selection problems with multiple, conflicting objectives and both qualitative and quantitative criteria " International Journal of Production Economics vol. 86 (3), Dec 2003. pp. 187-1999
31. G. Grau, J. P. Carvallo, X. Franch, and C. Quer, "DesCOTS: ASoftware System for Selecting COTS Components," inEUROMICRO'04, Rennes, France, 2004, pp. 118-126.
32. A. Mohamed, G. Ruhe, and A. Eberlein, "Decision Support for Handling Mismatches between COTS Products and System Requirements," in ICCBSS'07, Banff, Canada, 2007.
33. J. Kontio, OTSO: A Systematic Process for Reusable Software Component Selection, Univ. Maryland report CS-TR-3478, UMIACS-TR-95-63, 1995.
34. L. Chung and K. Cooper, "COTS-Aware Requirements Engineering and Software Architecting", Proceedings of the 4th International Workshop on System/Software Architectures (IWSSA), 2004
35. J. Bhuta, B. Boehm, "A Method for Compatible COTS Component Selection", ICCBSS, LNCS vol. 3412, Springer, 2005.
36. P. Avgeriou and N. Guelfi, "Resolving Architectural Mismatches of COTS through Architectural Reconciliation" in the 4th International Conference on COTS-BasedSoftware Systems (ICCBSS'05), Bilbao, Spain, 2005, pp. 248-257.
37. M. Shaw, "Architectural Issues in Software Reuse: It's Not Just the Functionality, It's the Packaging," in ACM SIGSOFT Symposium on Software Reusability, 1995, pp. 3-6.
38. Navarrete, Catalunya "How agile COTS selection methods are (and can be)?" Software Engineering and Advanced Applications, 2005. 31st EUROMICRO Conference on 30 Aug.-3 Sept. 2005 PP- 160 – 167
39. Ibrahim, Abdel-Halim, H. Far, Eberlein3 "UnHOS: A Method for Uncertainty Handling in Commercial Off-The-Shelf (COTS) Selection: International Journal of Energy, Information and Communications Vol. 2, Issue 3, August 2011

This page is intentionally left blank

Global Journal of Computer Science and Technology
Neural & Artificial Intelligence
Volume 12 Issue 10 Version 1.0 Year 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Relevance Search via Bipolar Label Diffusion on Bipartite Graphs

By Zhang Liang & Ren Lixiao
Tianjin University of Science and Technology Tianjin, China

Abstract - The task of relevance search is to find relevant items to some given queries, which can be viewed either as an information retrieval problem or as a semi-supervised learning problem. In order to combine both of their advantages, we develop a new relevance search method using label diffusion on bipartite graphs. And we propose a heat diffusion-based algorithm, namely bipartite label diffusion (BLD). Our method yields encouraging experimental results on a number of relevance search problems.

Keywords : *Relevance search, ranking, graph diffusion, bipartite graphs.*

GJCST-E Classification: *H.3.3*

Strictly as per the compliance and regulations of:

Relevance Search via Bipolar Label Diffusion on Bipartite Graphs

Zhang Liang [α] & Ren Lixiao [α]

Abstract -The task of relevance search is to find relevant items to some given queries, which can be viewed either as an information retrieval problem or as a semi-supervised learning problem. In order to combine both of their advantages, we develop a new relevance search method using label diffusion on bipartite graphs. And we propose a heat diffusion-based algorithm, namely bipartite label diffusion (BLD). Our method yields encouraging experimental results on a number of relevance search problems.

Keywords : Relevance search, ranking, graph diffusion, bipartite graphs.

I. Introduction

The problem of relevance search (RS) has been recognized as an important and interesting problem in machine learning and information retrieval, which refers to finding relevant items to a small query set given by the user. Along with the ability to find a cluster of data that shares some common properties to the query, it is also a primary goal of the information retrieval (IR) systems. Typical applications of RS include the discovery of relevant words in a text corpus[1-4], answers to community questions[5,6], features in concept-learning problems[7], recommendations in collaborative filtering systems[8,9], among others.

Google Sets [10] is a well-known and successful representative of RS systems that has been widely used. It uses vast amount of web-pages to create a list of related items from a few examples, such as people, movies, words, places, etc. Most of the other RS systems are similar to Google Sets in that they perform some ranking algorithm on a large corpus of documents or web-pages. Ghahramani and Heller [11] proposed the Bayesian Sets algorithm that uses a model-based concept of clusters and performs Bayesian inference to compute the ranking scores of items. Sun et al. [12] used bipartite graphs to model the data and introduced random walks with restarts to rank the items.

The RS problem can be viewed from several angles. First, finding relevant items to the query is a standard IR task[11,13], which may be solved using classic IR algorithms, such as HITS[14], nearest neighbors, naïve Bayes and Rocchio's algorithm. These algorithms can compute a list of items ordered by the relevance to the query. However, there is no explicit boundary between "relevant" and "irrelevant" items. Second, RS can be interpreted as a special case of semi-supervised learning (SSL) problem with a few positive examples given as the query[11]. It is essentially a one-class classification/clustering problem, which can be solved using one-class learning techniques like mapping-convergence [15] and OC-SVMs [16, 17]. The relevant and irrelevant items can be explicitly classified through SSL algorithms. However, most of these algorithms do not provide the rank order of items, which is of importance in IR systems.

In this paper, we propose and evaluate a new RS method called bipartite label diffusion (BLD) that can be seen as a hybrid between IR and SSL. BLD is a diffusion-based algorithm that works on bipartite graph. User queries are mapped to vertices with positive labels on the graph. BLD performs local computation at every vertex of the graph and develops a global classifier through the label diffusion process. The diffusion process is often modeled using a Markov chain on the graph. In BLD, we modified the diffusion model by adopting the "label spreading" method [18] to allow negative labels diffuse on the graph. If the word/document is relevant to the queries, the score value is positive, otherwise the value is negative. The more the item is relevant, the higher the score is. Thus BLD can be seen as a semi-supervised learning method.

The rest of this paper is organized as follows. We first model the corpus data as document-word bipartite graphs in section 2. The detail of the BLD algorithm is given in section 3. In section 4, we present empirical results to demonstrate the effectiveness of BLD on relevance search problems. Finally, some concluding remarks are given in section 5.

II. Bipartite Graph Model

Bipartite graph models have been successfully applied in many fields such as text clustering [19-21], collaborative filtering [22, 23], and content-based image retrieval [24]. The first step of our method is to model the document-word dataset as a bipartite model. We also employ a query expansion scheme to enhance the searching capacity. Detailed description is given in this section.

Suppose we are given a set of n documents $\mathrm{D} = \{d_1, ..., d_n\}$, which contain a set of m words $\mathrm{W} = \{w_1, ..., w_m\}$, the document-word collection can be modeled as a undirected bipartite graph $\mathrm{G} = (\mathrm{D} \cup \mathrm{W}, E)$

Author α : Food Safety Management and Strategy Research Center, School of Economics & Management, Tianjin University of Science and Technology Tianjin, China.

,where $E=\{(d_i,w_j): d_i \in \mathsf{D}, w_j \in \mathsf{W}\}$ is the set of edges between two sets of vertices D and W. In the bipartite graph G, there are no edges between words or between documents. An edge $e(i, j)$ exists if word w_j appears in document d_i. which is denoted by $w_j \sim d_i$. And it indicates an association between a word and a document, which may be quantitatively expressed by assigning positive weights on the edges. Assume we are given the user's query set of words W_Q ($\mathsf{W}_Q \neq \varnothing, \mathsf{W}_Q \subset \mathsf{W}$), or documents D_Q ($\mathsf{D}_Q \neq \varnothing, \mathsf{D}_Q \subset \mathsf{W}$), or both of them, the RS problem can be cast to a semi-supervised learning task where the query set contains the initial labeled examples. Fig. 1 shows a simple example of the bipartite graph model.

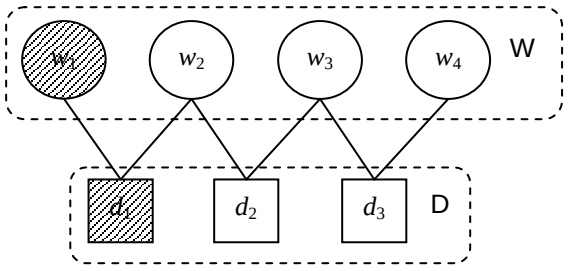


Figure 1: Example of a document-word bipartite graph which has 3 documents and a vocabulary of 4 words. The user's query set is $\mathsf{W}_Q=\{w_1\}$ and $\mathsf{D}_Q=\{d_1\}$. The goal of this similar sets retrieval problem is to discover a set of words that are similar to W_Q, and a set of documents that are similar to D_Q

There are many edge-weighting schemes in information retrieval research, among them we adopt the popular "tf-idf" weight:

$$\text{tf-idf}(d_i,w_j)=\frac{\ln \#(d_i,w_j)}{\ln \sum_k \#(d_i,w_k)} \cdot \ln \frac{n}{\left|\{d: w_j \sim d\}\right|}$$

where $\#(d_i,w_j)$ is the number of occurrence of word w_j in document d_i. Words with high tf-idf values imply a strong relationship with the document they appear in.
We can define an $n \times m$ edge weight matrix M:

$$M_{i,j}=\begin{cases}\text{tf-idf}(d_i,w_j), & \text{if edge } (d_i,w_j) \text{ exists},\\ 0, & \text{otherwise.}\end{cases}$$

To interpret the edge weight matrix from the perspective of Markov chains on graph, which will be described in section 3, we define the transition probability from d_i to w_j as:

$$P_{i,j}=M_{i,j}\Big/\sum_{p=1}^{m} M_{i,p} \qquad (1)$$

And the transition probability from w_j to d_i is given by,

$$Q_{j,i}=M_{i,j}\Big/\sum_{p=1}^{n} M_{p,j} \qquad (2)$$

The Markov transition matrix P normalizes M such that every row sum up to 1, and Q normalizes M^T such that every row sum up to 1.

Thus the $(m+n)\times(m+n)$ adjacency matrix A of the bipartite graph may be written as:

$$\mathbf{A}=\begin{pmatrix}\mathbf{0} & \mathbf{Q}\\ \mathbf{P} & \mathbf{0}\end{pmatrix},$$

Where the vertices of the graph is ordered such that the first m vertices index the words and the last n vertices index the documents.

Note that bipartite graph can also be a good model in the scenario of collaborative filtering or recommendation systems. A standard collaborative filtering problem often involves a user set U, an item set I, and a vote set V. In a similar manner, we can build a bipartite graph $\mathsf{G}=(\mathsf{U},\mathsf{I},E)$. The normalized vote scores of V can directly serve as the weight matrix. Moreover, the aim of collaborative filtering is to discover new items that an active user may like based on her voted items, which could be easily translated to a RS problem by finding similar items to the user's favorite ones (items with high voting scores).

III. Label Diffusion on Bipartite Graphs

In general, there are two primary approaches to cluster vertices on a bipartite graph: one is to partition the graph to disjoint parts corresponding to different clusters[25,26]; the other is to compute a rank value for each vertex indicating the probability that the vertex is in a cluster[24,27]. To capture the uncertainties on data-cluster assignments, we investigate the problem from the latter angle.

a) Markov Chains on the Graph

In a bipartite graph, there are no direct relationships among the words or among the documents. However, in diffusion-based methods[18,27], the strength of the similarities among elements on the same side of the bipartite graph may be captured by a local probability evolution process, which can be integrated to obtain a global view of the data.

We define a Markov chain to describe the diffusion process over the bipartite graph. Assume there is a discrete time random walk on the bipolar graph G, the row-normalized adjacency matrix $\mathbf{A}$ is the one-step transition matrix that $A_{i,j}$ is the probability of moving to the jth vertex of v given that the current step is at the ith vertex. More specifically, $\mathbf{P}$ is the document-word transition matrix containing the transition probabilities of

moving from a document vertex to a word vertex, and $\mathbf{Q}$ is the word-document transition matrix.

Since the document-word graph is bipartite, all the paths from w_i to w_j must go through vertices in D. As is shown in [15, 18], the similarity of two words w_i and w_j, is a quantity proportional to the probability of direct transitions between them, denoted by $p(w_i, w_j)$, $p(w_i, w_j) = p(w_i)p(w_j \mid w_i)$, where $p(w_i) = \deg(w_i)$ the degree of w_i, and $p(w_j \mid w_i)$ is the conditional transition probability from w_i to w_j,

$$p(w_j \mid w_i) = \sum_p p(d_p \mid w_i)p(w_j \mid d_p) = \sum_p Q_{i,p}P_{p,j}$$

Obviously, it is a 2-step stochastic process that first maps w_i to the document set, and then maps the documents back to w_j.

Similarly, the conditional transition probability from d_i to d_j is given by,

$$p(d_j \mid d_i) = \sum_p p(w_p \mid d_i)p(d_j \mid w_p) = \sum_p P_{i,p}Q_{p,j}$$

Therefore, by formulating the relationship between documents and words as Markov chains on the bipartite graph, we have the word-word transition probability matrix $\mathbf{QP}$, and the document-document transition probability matrix $\mathbf{PQ}$.

b) Diffusion Process

Intuitively, our bipolar diffusion framework works as following: First, we construct a bipartite graph with two poles as is described in section 2.3. The heat pole stands for the words and documents that are most relevant to the query, while the cold pole contains the "strongest negative" words and documents. Then a certain amount of heat is injected to the graph through the heat pole, and the cold polar extracts the heat out of the system as a heat sink. And the heat diffuses through the edges of the graph. Since the system has two poles, we name the heat diffusion as the "bipolar diffusion process".

The diffusion process may be thought of as a Markov chain on the graph. The fundamental property of a Markov chain is the Markov property, which make it possible to predict the future state of a system from its present state ignoring its past states. We denote $\boldsymbol{p}^{(t)}$ and $\boldsymbol{q}^{(t)}$ as the labels of documents and words at time *t*. The Markov process then defines a dynamic system,

$$\boldsymbol{p}^{(t+1)} = \mathbf{Q} \cdot \boldsymbol{q}^{(t)} = \mathbf{QP}\boldsymbol{p}^{(t)} \quad (3)$$

$$\boldsymbol{q}^{(t+1)} = \mathbf{P} \cdot \boldsymbol{p}^{(t)} = \mathbf{PQ}\boldsymbol{q}^{(t)} \quad (4)$$

This simple 2-step diffusion process captures the interaction between the two sets of vertices on the graph. It requires $\boldsymbol{p}^{(t)}$ and $\boldsymbol{q}^{(t)}$ be the probability distributions of Markov states with non-negative values. However, in our bipolar graph diffusion framework, the vertices of the cold pole are labeled by negative values. In the following, we cast the diffusion process as a semi-supervised learning problem, which makes it possible to diffuse heat and cold simultaneously on the graph.

First, we use an $(n+2)$-vector $\boldsymbol{p}^{(0)}$ to denote the initial labels of documents, $\boldsymbol{p}^{(0)} = \{1, \underbrace{0, \ldots 0}_{n}, -1\}^T$; and an $(m+2)$-vector $\boldsymbol{q}^{(0)}$ to denote the initial labels of words, $\boldsymbol{q}^{(0)} = \{1, \underbrace{0, \ldots 0}_{m}, -1\}^T$. The virtual vertices of heat pole are labeled by positive values, which allow heat diffuses through the graph; while negative values representing "cold" are assigned to the virtual vertices of cold pole. And the vertices keep exchanging these two kinds of energy as the diffusion process proceeds.

Then we adopt the "label spreading" method[18] to allow negative labels diffuse on the graph. The iteration in Eq. (3) and Eq. (4) can be rewritten as,

$$\boldsymbol{p}^{(t+1)} = \alpha \boldsymbol{p}^{(0)} + (1-\alpha)\mathbf{Q} \cdot \boldsymbol{q}^{(t)}$$

$$\boldsymbol{q}^{(t+1)} = \beta \boldsymbol{q}^{(0)} + (1-\beta)\mathbf{P} \cdot \boldsymbol{p}^{(t)}$$

Where α and β are scaling parameters both in (0,1), which specify the relative amount of the heat/cold a vertex received from its neighbors and its initial label information. To simplify the diffusion process, we set $\alpha = \beta = 1/2$.

The iteration equations indicate that when $t \geq 1$

$$\begin{aligned}\boldsymbol{p}^{(t+1)} &= \frac{1}{2}(\boldsymbol{p}^{(0)} + \mathbf{Q} \cdot \boldsymbol{q}^{(t)}) = \frac{1}{2}\boldsymbol{p}^{(0)} + \frac{1}{4}\mathbf{Q} \cdot \boldsymbol{q}^{(0)} + \frac{1}{4}\mathbf{QP} \cdot \boldsymbol{p}^{(t-1)} \\ &= (\sum_{i=0}^{t+1} \mathbf{\Lambda}^i)(\frac{1}{2}\boldsymbol{p}^{(0)} + \frac{1}{4}\mathbf{Q}\boldsymbol{q}^{(0)}) + \mathbf{\Lambda}^t \cdot \boldsymbol{p}^{(0)}\end{aligned} \quad (5)$$

$$\begin{aligned}\boldsymbol{q}^{(t+1)} &= \frac{1}{2}(\boldsymbol{q}^{(0)} + \mathbf{P} \cdot \boldsymbol{p}^{(t)}) = \frac{1}{2}\boldsymbol{q}^{(0)} + \frac{1}{4}\mathbf{P} \cdot \boldsymbol{p}^{(0)} + \frac{1}{4}\mathbf{PQ} \cdot \boldsymbol{q}^{(t-1)} \\ &= (\sum_{i=0}^{t+1} \mathbf{\Gamma}^i)(\frac{1}{2}\boldsymbol{q}^{(0)} + \frac{1}{4}\mathbf{P}\boldsymbol{p}^{(0)}) + \mathbf{\Gamma}^t \boldsymbol{q}^{(0)}\end{aligned} \quad (6)$$

Where $\mathbf{\Lambda} = \frac{1}{4}\tilde{\mathbf{Q}}\tilde{\mathbf{P}}$ and $\mathbf{\Gamma} = \frac{1}{4}\tilde{\mathbf{P}}\tilde{\mathbf{Q}}$. Since $\tilde{\mathbf{P}}$ and $\tilde{\mathbf{Q}}$ are row-normalized matrices, Eq. (5) follows that when $t \to \infty$, $\boldsymbol{p}^{(t+1)}$ converges to,

$$\boldsymbol{p}^{(\infty)} = (\mathbf{I} - \mathbf{\Lambda})^{-1}(\frac{1}{2}\boldsymbol{p}^{(0)} + \frac{1}{4}\mathbf{Q}\boldsymbol{q}^{(0)}) \quad (7)$$

And Eq. (6) converges to,

$$\boldsymbol{q}^{(\infty)} = (\mathbf{I} - \mathbf{\Gamma})^{-1}(\frac{1}{2}\boldsymbol{q}^{(0)} + \frac{1}{4}\mathbf{P}\boldsymbol{p}^{(0)}) \quad (8)$$

IV. Experiments

Our proposed BLD method is applicable to discover similar items to a query set from a corpus of text data or user rating data. In this section, we experiment with our model on a set of RS problems.

The experiments are performed on two standard text datasets: Reuters-21578[1], and a collaborative filtering dataset: MovieLens[2]. All of the text datasets were preprocessed by removing the stopwords and stemming. And for Reuters-21578, we use a subset of the ten most frequent categories with highest number of positive examples, namely Reuters-10. The main features of the datasets are summarized in Table 1.

Table I : Datasets Summary

Datasets	# documents/uses	# words/items	# nonzero entries
Reuters-10	9,989	5,180	373,440
MovieLens	943	1,682	100,000

We conducted two experiments on both the text dataset Reuter-10 and the movie rating dataset MovieLens to evaluate the RS ability of our method. The results and comparisons with Google Sets[3], Internet movie database (IMDB)[4], and Bayesian Sets are given in tables 2, 3 and 4. Concerning the application of movie recommending, the query takes the form of a set of movies. We regard movies as documents and users as words, and the rating scores are equivalent to word frequencies. Thus our algorithm can be easily adapted to collaborative filtering datasets.

Table II : Relevant Words Discovered By Google Sets and Bld Based on the Same Given Queries. Bld Runs on Reuters-10

query: science + technology		query: market + price	
Google Sets	BLD	Google Sets	BLD
science	scienc	market	market
technology	technologi	price	pric
business	univers	overview	money
sports	engineer	view	stock
health	educat	risk	valu
entertainment	research	gains	bond
education	comput	forecasts	busi
politics	life	war	compani
travel	develop	Intel	trade
computers	cultur	losses	economic

Table III : Relevant Movies Discovered By Google Sets and Bld Based on Query Full Metal Jacket. Bayesian Sets and Bld Run on Movielens

Query: Full Metal Jacket (1987)			
Google Sets	IMDB	Bayesian Sets	BLD
Saving Private Ryan (1998)	Platoon (1986)	The Terminator (1984)	Platoon (1986)
Apocalypse Now (1979)	All Quiet on the Western Front (1930)	Star Wars Episode V (1980)	Rambo:First Blood (1982)
Platoon (1986)	Cidade de Deus (2002)	Raiders of the Lost Ark (1981)	Braveheart (1995)
Pulp Fiction (1994)	Batoru Rowaiaru (2000)	Aliens (1986)	Apocalypse Now (1979)
Hamburger Hill (1987)	If… (1968)	Die Hard (1988)	Star Wars Episode V (1980)

Table IV : Relevant Movies Discovered By Google Sets and Bld Based on Query the Graduate. Bayesian Sets and Bld Run on Movielens

Query: The Graduate (1967)			
Google Sets	IMDB	Bayesian Sets	BLD
Chinatown (1974)	Mysterious Skin (2004)	Casablanca (1942)	Casablanca (1942)
Midnight Cowboy (1969)	Giant (1956)	The Wizard of Oz (1939)	Annie Hall (1977)
Annie Hall (1977)	The Notebook (2004)	One Flew over the Cuckoo's Nest (1975)	Gone with the Wind (1939)
Taxi Driver (1976)	Bigfish (2003)	The Godfather (1972)	To Kill a Mockingbird (1962)
Bonnie and Clyde (1967)	Notes on a Scandal (2006)	Amadeus (1984)	Giant (1956)

Among the tested algorithms and systems, Google Sets is a baseline RS system that is based on vast amount of web data. Although the Google Sets algorithm is not available for us to run it on our datasets, it is still informative and worth to be compared with; IMDB provides movie recommendations by generating a list of 5 movies most related to the query, which is based on collaborative filtering technology; Bayesian Sets views RS as a Bayesian inference problem and gives the corresponding ranking algorithm. We experiment with an online Bayesian Sets recommending system on Movie Lens dataset.

From the query results of the relevant words discovery task and the movie suggestion task, we can make the following comments:

1. Table 2 shows that both Google Sets and our method can achieve to some extent similar sets to the query. There is not an objective standard to tell the exact similarity between words or movies.

[1] http://www.daviddlewis.com/resources/testcollections/reuters21578/
[2] http://www.grouplens.org/system/files/ml-data.tar__0.gz
[3] http://labs.google.com/sets
[4] http://www.imdb.com/

However, the words that our method retrieved are obviously more sensible than some results of Google Sets, e.g. "war" and "Intel" for the query of "market" and "price". We think this is because Google Sets and our method have different learning mechanisms and are based on different corpuses.

2. Our method serves as a good algorithm for recommending systems on the MovieLens dataset. The recommended movies to the query "Full Metal Jacket" are all related to war. And for "The Graduate", most of the results are romance and drama movies. We notice that IMDB's suggestions are often popular and new, while Bayesian Sets and our method, limited by the MovieLens dataset, tend to generate classic movies.

V. Conclusion

In this paper we developed a new graph diffusion algorithm for RS. We used bipartite graphs to model the relationships between documents and words. We also modeled the diffusion process using a Markov chain on the graph, and presented the corresponding label diffusion algorithm. In future work, we will extend the proposed method to other applications (e.g. social networks, question answering systems).

VI. Acknowledgment

This work is supported by MOE (Ministry of Education in China) Liberal arts and Social Sciences Foundation under Grant No. 12YJC860056, and the Program for the Liberal arts and Social Sciences Research of Higher Learning Institutions of Tianjin under grant No. 20112107.

References Références Referencias

1. Blondel, V. D., Gajardo, A., Heymans, M., Senellart, P. and Dooren, P. V., 2004. A measure of similarity between graph vertices: Applications to synonym extraction and web searching. Siam Review. 46, (4), 647-666.
2. Lavrenko, V., 2004. A generative theory of relevance. University of Massachusetts Amherst.
3. Lin, D., 1998. Automatic retrieval and clustering of similar words. In: Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics (COLING/ACL-98), 768-774.
4. Wang, R. C. and Cohen, W. W., 2007. Language-independent set expansion of named entities using the web. In: Proceedings of the 17th IEEE International Conference on Data Mining, 342-350.
5. Deng, H., Lyu, M. R. and King, I., 2009. A generalized Co-HITS algorithm and its application to bipartite graphs. In: Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining, 239-248.
6. Wang, J., Robertson, S., Vries, A. P. and Reinders, M. J., 2008. Probabilistic relevance ranking for collaborative filtering. Inf. Retr. 11, (6), 477-497.
7. Cohen, W. W., 2000. Automatically extracting features for concept learning from the web. In: Proceedings of the 17th International Conference on Machine Learning, 159-166.
8. Basu, C., Hirsh, H. and Cohen, W., 1998. Recommendation as classification: using social and content-based information in recommendation. In: Proceedings of the 15th national/tenth conference on Artificial intelligence/Innovative applications of artificial intelligence, 714-720.
9. Wang, H. and Hancock, E. R., 2008. Probabilistic relaxation labelling using the Fokker-Planck equation. Pattern Recogn. 41, (11), 3393-3411
10. Google, Google Sets. <http://labs.google.com/sets>.
11. Ghahramani, Z. and Heller, K. A., 2005. Bayesian sets. In: Advances in Neural Information Processing Systems, 18,
12. Sun, J., Qu, H., Chakrabarti, D. and Faloutsos, C., 2005. Relevance search and anomaly detection in bipartite graphs. SIGKDD Explor. Newsl. 7, (2), 48-55.
13. Lafferty, J. and Zhai, C., 2002. Probabilistic relevance models based on document and query generation. In Language Modeling and Information Retrieval, Kluwer International Series on Information Retrieval.
14. Kleinberg, J. M., 1999. Authoritative sources in a hyperlinked environment. J. ACM. 46, (5), 604-632.
15. Yu, H., Han, J. and Chang, K. C.-C., 2004. PEBL: Web Page Classification without Negative Examples. IEEE Trans. on Knowl. and Data Eng. 16, (1), 70-81.
16. Manevitz, L. M. and Yousef, M., 2002. One-class svms for document classification. J. Mach. Learn. Res. 2, 139-154.
17. Schölkopf, B., Platt, J. C., Shawe-Taylor, J. C., Smola, A. J. and Williamson, R. C., 2001. Estimating the Support of a High-Dimensional Distribution. Neural Comput. 13, (7), 1443-1471.
18. Zhou, D., Bousquet, O., Lal, T. N., Weston, J. and Schölkopf, B., 2004. Learning with local and global consistency. In: Advances in Neural Information Processing Systems, 16, 595-602.
19. Dhillon, I. S., 2001. Co-clustering documents and words using bipartite spectral graph partitioning. In: Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, 269-274.
20. Hu, T., Qu, C., Tan, C. L., Sung, S. Y. and Zhou, W., 2006. Preserving Patterns in Bipartite Graph Partitioning. In: Proceedings of the 18th IEEE

International Conference on Tools with Artificial Intelligence, 489-496.
21. Sindhwani, V. and Melville, P., 2008. Document-Word Co-regularization for Semi-supervised Sentiment Analysis. In: Proceedings of the 2008 Eighth IEEE International Conference on Data Mining, 1025-1030.
22. Kunegis, J. and Lommatzsch, A., 2009. Learning spectral graph transformations for link prediction. In: Proceedings of the 26th Annual International Conference on Machine Learning, 561-568.
23. Shieh, J.-R., Yeh, Y.-T., Lin, C.-H., Lin, C.-Y. and Wu, J.-L., 2008. Collaborative knowledge semantic graph image search. In: Proceedings of the 17th international conference on World Wide Web, 1055-1056.
24. Bauckhage, C., 2007. Distance-free image retrieval based on stochastic diffusion over bipartite graphs. In: Proceedings of the 18th British Machine Vision Conference, 590-599.
25. Fern, X. Z. and Brodley, C. E., 2004. Solving cluster ensemble problems by bipartite graph partitioning. In: Proceedings of the 21st International Conference on Machine learning (ICML04), 36.
26. Rege, M., Dong, M. and Fotouhi, F., 2008. Bipartite isoperimetric graph partitioning for data co-clustering. Data Min. Knowl. Discov. 16, (3), 276-312.
27. Yu, K., Yu, S. and Tresp, V., 2005. Soft clustering on graphs. In: Advances in Neural Information Processing Systems 18.

Global Journals Inc. (US) Guidelines Handbook 2012

www.GlobalJournals.org

FELLOW OF ASSOCIATION OF RESEARCH SOCIETY IN COMPUTING (FARSC)

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

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

MEMBER OF ASSOCIATION OF RESEARCH SOCIETY IN COMPUTING (MARSC)

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

AUXILIARY MEMBERSHIPS

ANNUAL MEMBER

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

(II) Choose corresponding Journal.

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

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

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

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

Preferred Author Guidelines

MANUSCRIPT STYLE INSTRUCTION (Must be strictly followed)

Page Size: 8.27" X 11'"

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

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

1. General,

2. Ethical Guidelines,

3. Submission of Manuscripts,

4. Manuscript's Category,

5. Structure and Format of Manuscript,

6. After Acceptance.

1. GENERAL

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

Scope

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

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

2. ETHICAL GUIDELINES

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

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

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

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

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

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

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

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

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

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

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

Please mention proper reference and appropriate acknowledgements wherever expected.

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

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

3. SUBMISSION OF MANUSCRIPTS

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

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

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

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

4. MANUSCRIPT'S CATEGORY

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

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

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

Research articles: These are handled with small investigation and applications

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

5.STRUCTURE AND FORMAT OF MANUSCRIPT

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

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

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

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

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

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

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

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

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

(h) Brief Acknowledgements.

(i) References in the proper form.

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

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

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

Format

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

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

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

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

Structure

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

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

Abstract, used in Original Papers and Reviews:

Optimizing Abstract for Search Engines

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

Key Words

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

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

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

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

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

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

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

Acknowledgements: Please make these as concise as possible.

References

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

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

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

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

Tables, Figures and Figure Legends

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

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

Preparation of Electronic Figures for Publication

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

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

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

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

6. AFTER ACCEPTANCE

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

6.1 Proof Corrections

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

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

(Free of charge) from the following website:

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

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

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

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

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

6.3 Author Services

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

6.4 Author Material Archive Policy

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

6.5 Offprint and Extra Copies

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

the search? Will I be able to find all information in this field area? If the answer of these types of questions will be "Yes" then you can choose that topic. In most of the cases, you may have to conduct the surveys and have to visit several places because this field is related to Computer Science and Information Technology. Also, you may have to do a lot of work to find all rise and falls regarding the various data of that subject. Sometimes, detailed information plays a vital role, instead of short information.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

31. Adding unnecessary information: Do not add unnecessary information, like, I have used MS Excel to draw graph. Do not add irrelevant and inappropriate material. These all will create superfluous. Foreign terminology and phrases are not apropos. One should NEVER take a broad view. Analogy in script is like feathers on a snake. Not at all use a large word when a very small one would be

sufficient. Use words properly, regardless of how others use them. Remove quotations. Puns are for kids, not grunt readers. Amplification is a billion times of inferior quality than sarcasm.

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

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

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

Informal Guidelines of Research Paper Writing

Key points to remember:

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

Final Points:

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

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

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

General style:

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

To make a paper clear

· Adhere to recommended page limits

Mistakes to evade

- Insertion a title at the foot of a page with the subsequent text on the next page

- Separating a table/chart or figure - impound each figure/table to a single page
- Submitting a manuscript with pages out of sequence

In every sections of your document

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

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

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

· Align the primary line of each section

· Present your points in sound order

· Use present tense to report well accepted

· Use past tense to describe specific results

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

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

Title Page:

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

Abstract:

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

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

Write your summary when your paper is completed because how can you write the summary of anything which is not yet written? Wealth of terminology is very essential in abstract. Yet, use comprehensive sentences and do not let go readability for briefness. You can maintain it succinct by phrasing sentences so that they provide more than lone rationale. The author can at this moment go straight to

shortening the outcome. Sum up the study, with the subsequent elements in any summary. Try to maintain the initial two items to no more than one ruling each.

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

Approach:

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

Introduction:

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

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

Approach:

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

Procedures (Methods and Materials):

This part is supposed to be the easiest to carve if you have good skills. A sound written Procedures segment allows a capable scientist to replacement your results. Present precise information about your supplies. The suppliers and clarity of reagents can be helpful bits of information. Present methods in sequential order but linked methodologies can be grouped as a segment. Be concise when relating the protocols. Attempt for the least amount of information that would permit another capable scientist to spare your outcome but be cautious that vital information is integrated. The use of subheadings is suggested and ought to be synchronized with the results section. When a technique is used that has been well described in another object, mention the specific item describing a way but draw the basic

principle while stating the situation. The purpose is to text all particular resources and broad procedures, so that another person may use some or all of the methods in one more study or referee the scientific value of your work. It is not to be a step by step report of the whole thing you did, nor is a methods section a set of orders.

Materials:

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

Methods:

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

Approach:

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

What to keep away from

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

Results:

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

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

Content

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

What to stay away from

- Do not discuss or infer your outcome, report surroundings information, or try to explain anything.
- Not at all, take in raw data or intermediate calculations in a research manuscript.

- Do not present the similar data more than once.
- Manuscript should complement any figures or tables, not duplicate the identical information.
- Never confuse figures with tables - there is a difference.

Approach

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

Figures and tables

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

Discussion:

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

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

Approach:

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

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

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

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

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

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

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

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

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

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

Index

T

U

W

save our planet

Global Journal of Computer Science and Technology

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

ISSN 9754350