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

- Artificial Neural Network
- Estimation of Global Solar
- Data Mining Technique
- Intelligent Intrusion Detection

Space Justin

Volume 12 | Issue 11 | 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 11 (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 11 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

Recognition of Similar Shaped Handwritten Marathi Characters Using Artificial Neural Network

By Mrs. Archana P.Jane & Mukesh. A.Pund
Ram Meghe Institute of Research & Technology, Badnera

Abstract - The growing need have handwritten Marathi character recognition in Indian offices such as passport, railways etc has made it vital area of a research. Similar shape characters are more prone to misclassification. In this paper a novel method is provided to recognize handwritten Marathi characters based on their features extraction and adaptive smoothing technique. Feature selections methods avoid unnecessary patterns in an image whereas adaptive smoothing technique form smooth shape of charecters. Combination of both these approaches leads to the better results. Previous study shows that, no one technique achieves 100% accuracy in handwritten character recognition area. This approach of combining both adaptive smoothing & feature extraction gives better results (approximately 75-100) and expected outcomes.

Keywords : *Character recognition, features Extraction, adaptive Smoothing, Image segmentation, pattern matching, Image Pixel Rating.*

GJCST-E Classification: *I.5.2*

Strictly as per the compliance and regulations of:

Recognition of Similar Shaped Handwritten Marathi Characters Using Artificial Neural Network

Mrs. Archana P.Jane [α] & Mukesh. A.Pund [σ]

Abstract - The growing need have handwritten Marathi character recognition in Indian offices such as passport, railways etc has made it vital area of a research. Similar shape characters are more prone to misclassification. In this paper a novel method is provided to recognize handwritten Marathi characters based on their features extraction and adaptive smoothing technique. Feature selections methods avoid unnecessary patterns in an image whereas adaptive smoothing technique form smooth shape of charecters. Combination of both these approaches leads to the better results. Previous study shows that, no one technique achieves 100% accuracy in handwritten character recognition area. This approach of combining both adaptive smoothing & feature extraction gives better results (approximately 75-100) and expected outcomes.

Keywords : Character recognition, features Extraction, adaptive Smoothing, Image segmentation, pattern matching, Image Pixel Rating.

I. Introduction

In recent years, handwritten Marathi character recognition has grabbed a lot of attention as Marathi being primary official language in Maharashtra has wide application in areas like passport, railways, postal address reading etc. Handwritten character recognition [4] consist of six main steps.

i. Handwritten drawn of character
ii. Training on handwritten characters.
iii. Testing on handwritten characters.

- Pixel rating an Image
- Image smoothing
- Features extraction.
- Image segmentation.
- Patterns matching
- Result display as per their % of pattern matched.

Earlier, traditional classifiers such as Nearest Neighbor (NN), Hidden Markov Models (HMM) etc. were adopted for character recognition, however they exhibit certain limitations. Machine learning (ML) algorithms [6] provide a promising alternative in character recognition based on the feature set given to them. Each character image sample can be expressed in terms of some 26 Computer Science & Information Technology (CS & IT) quantifiable attributes called features. A variety of features can be extracted such as primitives, profiles etc. Multi Layer (ML) algorithm is then trained with this list of measured features, so that it maps these input features onto a class among certain predefined classes [1, 2]. Then the classifier can be used to determine the class of unknown samples used for testing.

II. Literature Survey

Character recognition task has been attempted through many different approaches like template matching, statistical techniques like NN, HMM, Quadratic Discriminant function (QDF) etc. Template matching works effectively for recognition of standard fonts, but gives poor performance with handwritten characters and when the size of dataset grows. It is not an effective technique if there is font discrepancy [4]. HMM models achieved great success in the field of speech recognition in past decades, however developing a 2-D HMM model for character recognition is found difficult and complex [5]. NN is found very computationally expensive in recognition purpose [6]. N. Araki et al. [7] applied Bayesian filters based on Bayes Theorem for handwritten character recognition. Later, discriminative classifiers such as Artificial Neural Network (ANN) and Support Vector Machine (SVM) grabbed a lot of attention. In [3] G. Vamvakas et al. compared the performance of three classifiers: Naive Bayes, K-NN and SVM and attained best performance with SVM. However SVM suffers from limitation of selection of kernel. ANNs can adapt to changes in the data and learn the characteristics of input signal [8].Also, ANNs consume less storage and computation than SVMs [9]. Mostly used classifiers based on ANN are MLP and RBFN. B.K. Verma [10] presented a system for HCR using MLP and RBFN networks in the task of handwritten Hindi character recognition. The error back propagation algorithm was used to train the MLP networks. J. Sutha et al. in [11] showed the effectiveness of MLP for Tamil HCR using the Fourier descriptor features. R. Gheroie et al. in [12] proposed handwritten Farsi character recognition using MLP trained with error back propagation algorithm. Computer Science & Information Technology (CS & IT) 27 similar

Author α : M.E (Scholar), Prof.Ram Meghe Institute of Research & Technology, Badnera. E-mail : apjane@rediffmail.com
Author σ : Associate Prof. Prof. Ram Meghe Institute of Research & Technology, Bandera. E-mail : pundmukesh@rediffmail.com

shaped characters are difficult to differentiate because of very minor variations in their structures. In [13] T. Wakabayashi et al. proposed an F-Ratio (Fisher Ratio) based feature extraction method to improve results of similar shaped characters. They considered pairs of similar shaped characters of different scripts like English, Arabic/Persian, Devnagri, etc. and used QDF for recognition purpose. QDF suffers from limitation of minimum required size of dataset. F. Yang et al. in [14] proposed a method that combines both structural and statistical features of characters for similar handwritten Chinese character recognition. As it can be seen that various feature extraction methods and classifiers have been used for character recognition by researchers that are suitable for their work, we propose a novel feature set that is expected to perform well for this application. In this paper, the features are extracted on the basis of character geometry, which are then fed to each of the selected ML algorithms for recognition of SSHMC.

III. Proposed Method

In this paper, we proposed a novel method based on combinations on pixel rating an image, feature extractions and Image pattern matching. Proposed method gives considerable expected outputs than previous proposed character recognition algorithms like HMM, NN, ML etc. Proposed method consists of following phases.

a) Training

In this proposed work, Training on images consists of listing all handwritten images with respect to its standard Marathi character images. See the training set data in the following figure.

Handwritten Images	Standard Marathi Character Images
अ	अ
आ	आ
इ	इ
ई	ई
उ	उ
स,	स,
क	क
र	र
न	न
म	म
द	द
प	प

In Training phase, we are having 51 Marathi characters and we Trained 20 handwritten characters with respect to each corresponding standard Marathi character images. The no. Of images in Training affects turnaround time of entire process execution.

$$Ti \propto N \text{ ----------Eq.3.1.1}$$

Where
T_i=Turnaround time.
N=No. of Training images.

An accuracy of result depends on the no. of trained handwritten images per standard Marathi character image.

$$Ac \propto \frac{Him}{Sim} \text{----------Eq.2.1.2}$$

Where
A_c= Accuracy of character Recognition
H_{im}=No. of handwritten patterns.
S_{im}= Standard Marathi image=1

b) Testing

Testing is a phase where no. of Marathi handwritten character image is tested against training set handwritten images.

Testing consists of following phases:

c) RGB-to-Binary image conversion

As we maintained Training images as a binary images, there is need to convert testing image into binary image. Binary image avoids unnecessary image segmentation and features extraction.

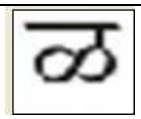
	
Input handwritten image	Input binary image

Table 3.2.2.1

d) Pixel rate an image

	
Input binary image	Pixel Rate Image

Table 3.2.2.2

Pixel rate an image used to identify an image pixel value either on or off. We set pixels height and width equal to 10 of which we got result shown in above table 3.2.2.2. Height, Width of an pixel rate image should be proportionate to size of an binary image.

From the table 3.2.2.2, it is observed that having pixels size 10 x 10 leads to loss of pixels which are represented with an equation

$$Pl = \int_1^{n} Pnhw(\log_2 n)$$----------Eq.3.2.1

Where
P_l= No. of Pixels loss
P_{nhw}= Pixels new height, width.
n=No. of off pixels in binary image.

e) Image edge smoothing

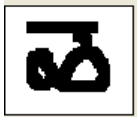
From literature review, it is observed that patterns in handwritten characters have large deviation factor with respect to its standard image pattern. Deviation factor is a mod difference value between no. of pixels patterns in training and testing images and represented with

$$Df = |Tr - Ts|$$-----eq.3.3.1

Where
D_f= Deviation factor.
T_r= No. of segments of training image.
T_s= No. of segments of testing image.

$$Tr = Ts = \sum_1^s \max_{1 \le hw \le 100} d\left(\frac{s}{\le n}\right)$$--------Eq.3.3.2

Where
S= size of an image.
hw= new height, width of segment.
D_n = new segment size

	
Input binary image	**Image After Edge Smoothing**

Average searching time for character recognition of test image uses best-fit search approach with training set images.

$$O(t) = O(N)$$-------eq.3.3.3

Where
O (t) =Average searching time.

f) Image segmentation

To recognize character, segmentation is done based on their patterns of size 2x2 as per equation 3.3.2.

g) Pattern Matching

We form fuzzy pattern match of pixels value (Equivalent decimal value) of segmented image based on which we match an object array patterns with following fuzzy rules:

$$Ptm = \sum_{i=1}^{\frac{s}{4}} Tsi \; v \left(Tri \; v \, Tri + 1 \, v \ldots\ldots v \, \frac{Trs}{4}\right)$$----eq3.5.1

Where
P_{tm}= fuzzy pattern match.
S= size of training image=size of testing image.

h) Experiment result

Sr No	Test Image	Recognize Image 1	Recognize Image 2	Recognize Image 3	Recognize Image 4	Recognize Image 5	Recognize Image 6	Recognize Image 7	Recognize Image 8	Recognize Image 9
1	प	प	व	ब	फ	म	क	ष	भ	य
Std.marathi char.		प	व	ब	फ	म	क	ष	भ	य
Deviation Factor		0.0000	0.1923	0.2345	0.2378	0.3124	0.3564	0.4219	0.5638	0.5698
Recognize Character								प		
Sr No	Test Image	Recognize Image 1	Recognize Image 2	Recognize Image 3	Recognize Image 4	Recognize Image 5	Recognize Image 6	Recognize Image 7	Recognize Image 8	Recognize Image 9
2	म	म	प	भ	फ	ष	स	य	क	न
Std.marathi char.		म	प	भ	फ	ष	स	य	क	न
Deviation Factor		0.0000	0.1276	0.2234	0.2567	0.3015	0.3567	0.4216	0.5216	0.5767
Recognize Character								म		
Sr no	Test Image	Recognize Image 1	Recognize Image 2	Recognize Image 3	Recognize Image 4	Recognize Image 5	Recognize Image 6	Recognize Image 7	Recognize Image 8	Recognize Image 9
3	फ	फ	प	व	क	अ	स	ष	थ	घ
Std.marathi char.		फ	प	व	क	अ	स	ष	थ	घ
Deviation Factor		0.0000	0.1435	0.24545	0.2535	0.25467	0.258965	0.46556	0.525464	0.58967
Recognize Character								फ		

Experiment Result 1

Sr.no.	No. of Training Images / standard Marathi characters	Type of Images	% Loss of Pixels/Image in Pixel rating Image	Average % Accuracy
1	10	JPG	33%	70%-80%
2	20	JPG	33%	80%-85%
3	30	JPG	33%	85%-90%
4	50	JPG	33%	90%-98%

Experiment Result 2

IV. Conclusion

From the study of Literature survey and proposed method, we conclude that, proposed method gives considerable and expected accuracy than previous character recognitions techniques like HMM, ML, NBP etc. Experiment results shows that, proposed method achieved an accuracy nearer to 98%provided no. of training samples per standard Marathi images should be maximum as possible as.

V. Future Work

In the process of recognizing handwritten character, human brains may fails that's why to keep an expectations to achieve 100% accuracy is not expectable. A future work is needed to correctly analyze segments patterns and fuzzy rules mentioned in an equation 3.5.1 to achieve better accuracy which should be independent of no. of training set images.

References Références Referencias

1. T.M. Mitchell, "Machine Learning", McGraw-Hill Education (ISE Editions), December 1997.
2. M. Zafar, D. Mohamad, M.M. Anwar, "Recognition of Online Isolated Handwritten Characters by Back propagation Neural Nets Using Sub-Character Primitive Features", IEEE Multitopic Conference (INMIC), 2006, pp. 157 – 162
3. G. Vamvakas, B. Gatos, S. Petridis, N. Stamatopoulos, "An Efficient Feature Extraction and Dimensionality Reduction Scheme for Isolated Greek Handwritten Character Recognition", IEEE Ninth International Conference on Document Analysis and Recognition(ICDAR), 2007, vol. 2, pp. 1073 – 1077 Figure 2. Misclassification Rate of Bayesian Network and C4.5 with FULL,CFS and CON Features Sets 34 Computer Science & Information Technology (CS & IT)
4. J.R. Prasad, U.V. Kulkarni, R.S. Prasad, "Offline handwritten character recognition of Gujrati script using pattern matching", IEEE 3rd International Conference on Anti-counterfeiting, Security, and Identification in Communication, 2009, pp. 611-615.
5. H.S. Park, S.W. Lee, "An HMMRF-Based Statistical Approach for Off-line Handwritten Character Recognition", IEEE Proceedings of the 13th International Conference on Pattern Recognition, 1996, vol. 2, pp. 320 – 324.
6. C.L. Liu, H. Sako, H. Fujisawa, "Performance evaluation of pattern classifiers for handwritten character recognition", International Journal on Document Analysis and Recognition (IJDAR), 2002, vol. 4, pp. 191–204.
7. N. Araki, M. Okuzaki, Y. Konishi , H. Ishigaki , "A Statistical Approach for Handwritten Character Recognition Using Bayesian Filter", IEEE 3rd International Conference on Innovative Computing Information and Control, 2008, pp. 194 – 194.
8. N. Arica, F.T. Yarman-Vural, "An Overview Of Character Recognition Focused On Off-line Handwriting", IEEE Transactions on Systems, Man, and Cybernetics, 2001, vol. 31, pp. 216 –233.
9. F. Kahraman, A. Capar, A. Ayvaci, H. Demirel, M. Gokmen, "Comparison of SVM and ANN performance for handwritten character classification", Proceedings of the IEEE 12th Signal Processing and Communications Applications Conference, 2004, pp. 615 – 618.
10. B.K. Verma, "Handwritten Hindi Character recognition Using Multilayer Perceptron and Radial Basis Function Neural Networks," Proceedings of IEEE International conference on Neural Networks, 1995, vol. 4, pp. 2111-2115.
11. J. Sutha, N. Ramaraj, "Neural Network Based Offline Tamil Handwritten Character Recognition System", IEEE International Conference on Computational Intelligence and Multimedia Applications, 2007, vol. 2, pp. 446 – 450.
12. R. Gharoie, M. Farajpoor, "Handwritten Farsi Character Recognition using Artificial Neural Network", International Journal of Computer Science and Information Security, 2009, vol. 4.
13. T. Wakabayashi, U. Pal, F. Kimura, Y. Miyake, "F-ratio Based Weighted Feature Extraction for Similar Shape Character Recognition" IEEE 10th International Conference on Document Analysis and Recognition (ICDAR), 2009, pp. 196-200.
14. F. Yang, X.D. Tian, X. Zhang, X.B. Jia, "An Improved Method for Similar Handwritten Chinese Character Recognition", IEEE Third International Symposium on Intelligent Information Technology and Security Informatics (IITSI), 2010, pp. 419 – 422.
15. N.J. Nilsson, "Introduction to Machine learning", An early draft of a proposed textbook, Robotics Lab, Deptt of Computer Science, Stanford University, Stanford, CA 94305, 1998.[16] WEKA 3: Data Mining With Open Source Machine Learning Software in JAVA URL: http://www.cs. waikato. ac. nz/ml/weka/.
16. N. Friedman, D. Geiger, M. Goldszmidt, "Bayesian Network Classifiers, Machine learning, 1997,pp. 131-163.

This page is intentionally left blank

Global Journal of Computer Science and Technology
Neural & Artificial Intelligence
Volume 12 Issue 11 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

Comparison of Angstrom-Prescott, Multiple Regression and Artificial Neural Network Models for the Estimation of Global Solar Radiation in Warri, Nigeria

By Ibeh G.F, Agbo G.A, Oboma D.N, Ekpe J.E & Odoh S
Ebonyi State University, Abakaliki

Abstract - In this paper, the application of artificial neural network, Angstrom-Prescott and multiple regressions models to study the estimation of global solar radiation in Warri, Nigeria for a time period of seventeen years were carried out. Our study based on Multi-Layer Perceptron (MLP) of artificial neural network was trained and tested using seventeen years (1991-2007) meteorological data. The error results and statistical analysis shows that MLP network has the minimum forecasting error and can be considered as a better model to estimate global solar radiation in Warri compare to the estimation from multiple regressions and Angstrom-Prescott models.

Keywords : Artificial neural networks, multiple regression, Angstrom-Prescott, prediction, troposphere, Multi-layer perceptron.

GJCST-E Classification : I.2.6

Strictly as per the compliance and regulations of:

Comparison of Angstrom-Prescott, Multiple Regression and Artificial Neural Network Models for the Estimation of Global Solar Radiation in Warri, Nigeria

Ibeh G.F [α], Agbo G.A [α], Oboma D.N [α], Ekpe J.E [α], & Odoh S [σ]

Abstract - In this paper, the application of artificial neural network, Angstrom-Prescott and multiple regressions models to study the estimation of global solar radiation in Warri, Nigeria for a time period of seventeen years were carried out. Our study based on Multi-Layer Perceptron (MLP) of artificial neural network was trained and tested using seventeen years (1991-2007) meteorological data. The error results and statistical analysis shows that MLP network has the minimum forecasting error and can be considered as a better model to estimate global solar radiation in Warri compare to the estimation from multiple regressions and Angstrom-Prescott models.

Keywords : Artificial neural networks, multiple regression, Angstrom-Prescott, prediction, troposphere, Multi-layer perceptron.

I. Introduction

The troposphere is the lower layer of the Earth's atmosphere. Most of the weather phenomena, systems, convection, turbulence and clouds occur in this layer, although some may extend into the lower portion of the stratosphere. The troposphere contains almost all the atmospheric water vapour. It contains about 70 to 80 percent of the total mass of the earth's atmosphere and 99 per cent of the water vapour. Temperature and water vapour content in the troposphere decrease rapidly with altitude and thus most of the water vapour in the troposphere is concentrated in the lower, warmer zone. Water vapor concentrations vary with latitude. The condition of the atmosphere as dictated by the sun is very dynamic both in space and time scales. The resulting solar interactions on the atmosphere leads to changes in weather as well as the so called climate change.

The objective of this study is to model the relevant data provided by the Nigerian Meteorological Agency, Federal Ministry of Aviation, Oshodi, Lagos, Nigeria as shown in Table 1 and the global solar radiation data collected from Renewable Energy for Rural Industrialization and Development in Nigeria using the Angstrom-Prescott, statistical technique (multiple regression model) and the artificial neural networks (ANN), and then comparing the results of these three models along with the measured solar radiation data. The table below (Table1) shows the atmospheric parameters in preparation of the prediction.

Table 1 : Meteorological Data for Warri

Month	$\bar{n}$ Hour	T_m °C	$\bar{H}_M$ (MJm^2day^{-1})
JAN	4.72	33.00	12.52
FEB	4.80	33.68	14.26
MAR	4.61	33.45	15.64.
APR	4.92	32.86	18.11
MAY	4.89	31.93	12.84
JUN	3.86	30.53	13.99
JUL	2.27	28.77	14.67
AUG	2.31	28.89	13.85
SEPT	2.57	29.99	15.40
OCT	4.15	31.28	16.55
NOV	5.23	32.74	15.81
DEC	5.66	32.66	18.16

Author α : Department of Industrial Physics, Ebonyi state University, Abakaliki Nigeria. E-mail : ibehgabriel@ymail.com

Author σ : Nigeria Turkish International College, Kano.

II. Methods and Procedures Multiple Regressions

Regression method is one of the most widely used statistical techniques Mendenhall and Beaver (1994) and Sharda, and Patil (1990). Multiple regression analysis is a multivariate statistical technique used to examine the relationship between a single dependent variable and a set of independent variables. The objective of the multiple regression analysis is to use independent variables whose values are known to predict the single dependent variable (Olaniyi, 2011). The effect of independent variables on the response is expressed mathematically by the regression or response function f:

$$y = f(x_1, x_2, \ldots, x_n; \beta_1, \beta_2, \ldots, \beta_n) \quad (1)$$

where y is the dependent variable.

$\beta_1, \beta_2, \ldots, \beta_n$ in the regression parameters which will be determined

The regression model for the observed response variable is written as

$$z = y + \varepsilon = f(x_1, x_2, \ldots, x_n; \beta_1, \beta_2, \ldots \beta_n) + \varepsilon \quad (2)$$

Where ε is the error in observed value z (Olaniyi, 2011).

To find *unknown regression parameters* $\{\beta_1, \beta_2, \ldots \beta_n\}$, *the method of least squares* (Beenstock, 1992) can be applied:

$$E(\beta_1, \beta_2, \ldots \beta_n) = \sum_{j=i}^{n} (z_j - y_j)^2 = \sum (z_j - f(x_1 x_2, \ldots, x_n; \beta_1, \beta_2, \ldots \beta_n))^2 \quad (3)$$

Where $E(\beta_1, \beta_2, \ldots \beta_n)$ is the error function or sum of squares of the deviations.

To estimate $\beta_1, \beta_2, \ldots n$ we minimize, E, by solving the system of equations:

$$\frac{\partial E}{\partial \beta_1} = 0, i = 1,2, \ldots, n \quad (4)$$

a) Model Specification and Analysis

The regression model to consider in this study takes the maximum temperature (Tm) and sunshine hour ($\bar{n}$) as the explanatory variables and measured values (H_m) as dependent variable. This is used to obtain a reliable parameter estimates in the regression. The model to be used can be specified as

$$H_M = f(T_M, \bar{n}) \quad (5)$$

More precisely;

$$H_M = \beta_0 + \beta_1 T_M + \beta_2 \bar{n} \quad (6)$$

$\beta_0, \beta_1, \beta_2 > 0$

III. Angstrom-Prescott Model

There are several types of empirical formulae for predicting the monthly mean daily global solar radiation as a function of readily measured climatic data (Iqbal, 1977; and Klien, 1977). Among the existing correlations, the one used in this paper is the Angstrom–Prescott regression equation, which relates the monthly mean daily global solar radiation to the number of hours of bright sunshine as follows

$$\frac{\bar{H}_M}{\bar{H}_O} = a + b\frac{\bar{n}}{\bar{N}} \quad (7)$$

Where,

$\bar{H}_M$ = measured monthly mean daily global solar radiation on a horizontal sunshine.

$\bar{n}$ = monthly mean daily bright sunshine hour, $\bar{N}$ = Maximum possible monthly mean daily sunshine,

a and b = regression constant.

The monthly mean daily extraterrestrial radiation $\bar{H}_o$ and monthly mean day length $\bar{N}$ was derived from the formulae:

$$\bar{H}_O = \frac{24}{\pi} I_{SC} E_O [W_S \sin\Phi \sin\Phi + \cos\Phi \cos\delta \cos W_S] \quad (8)$$

$$\overline{N} = \frac{2}{25} \text{COS-1 (- tan } \Phi \text{ tan } \delta) \quad (9)$$

$$\delta = 23.45 \sin\left[\frac{N+284}{365}\right] \quad (10)$$

$$W_S = \text{COS -1 (- tan } \Phi \text{ tan } \delta) \quad (11)$$

I_{SC} = solar constant (4.921 MJM^{-2} day^{-1})

N = characteristic day number

Φ = latitude angle

δ = angle of declination

$$a = -0.110 + 0.235 \cos\Phi + 0.323 \text{ (n/N)} \quad (12)$$

$$b = 1.449 - 0.553 \text{ Cos } \Phi - 0.694 \text{ (n/N)} \quad (13)$$

IV. Artificial Neural Network Model

Artificial neural network models are based on the neural structure of the brain (Ibeh et al, 2012). The brain learns from experience and so do artificial neural networks. Previous research has shown that artificial neural networks are suitable for pattern recognition, prediction and pattern classification tasks due to their nonlinear nonparametric adaptive-learning properties. As a useful analytical tool, ANN is widely applied in analyzing the data stored in database or data warehouse nowadays (Massie, 2001). One critical step in neural network application is network training. Generally, meteorological data is selected and refined to form training data sets. Artificial Neural Network is widely used in various branches of engineering and science and their unique proper of being able to approximate complex and nonlinear equations makes it a useful tool in quantitative analysis (Olaniyi, 2011).

The true power and advantage of neural networks lies in their ability to represent both linear and

non-linear relationships and in their being modeled. Traditional linear models are simply inadequate when it comes to modeling data that contains non-linear characteristics. In this paper, one model of neural network is selected among the main network architectures used. The basis of the model is neural structure as shown in Fig. 1. These neurons act like parallel processing units.

a) Multi-Layer Perception

The most popular network architecture in use today is perhaps the Multilayer Perceptron network and it uses supervised network (Hair and Tatham, 1998). This type of neural network is known as a supervised network because it requires a desired output in order to learn. The goal of this type of network is to create a model that correctly maps the input to the output using historical data so that the model can then be used to produce the output when the desired output is unknown. Figure 1 show the block diagram of a double hidden layer multiplayer perceptron (MLP).

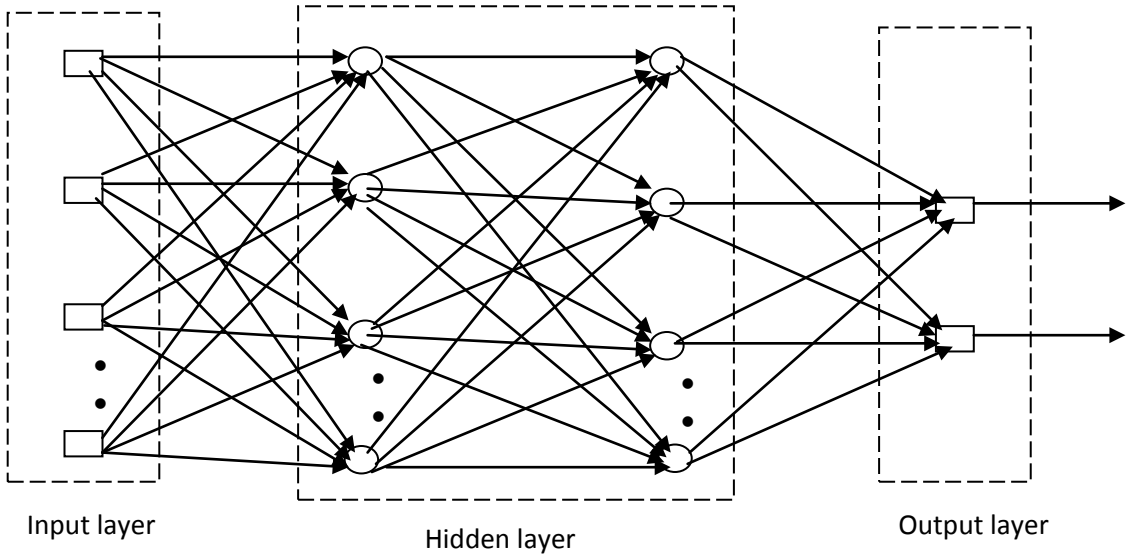


Figure 1: Structure of the artificial neural network with Double hidden layer

The inputs are fed into the input layer and get multiplied by interconnection weights as they are passed from the input layer to the hidden layer. Within the hidden layer, they get summed then processed by a nonlinear function (usually the hyperbolic tangent).

If more than a single hidden layer exists then, as the processed data leaves the first hidden layer, again it gets multiplied by interconnection weights, then summed and processed by the second hidden layer and so on. Finally the data is multiplied by interconnection weights then processed one last time within the output layer to produce the neural network output.

An artificial neuron is a unit that performs a simple mathematical operation on its inputs and imitates the function of biological neurons and their unique process of learning (Ibeh and Agbo, 2012). The weighted sum of the inputs are calculate using the following equation,

$$v_k = \sum_{j=1}^{m} x_j w_{kj} + b_k \qquad (7)$$

Where v_k is the weight sum from the *k*th hidden node, w_{kj} is the weight on connection from the *j*th to the *k*th node; x_j is an input data from input node; m is the total number of input (N=17); and b_k denotes a bias on the *k*th hidden node.

The mathematical structure of the normal method is as shown in fig.2

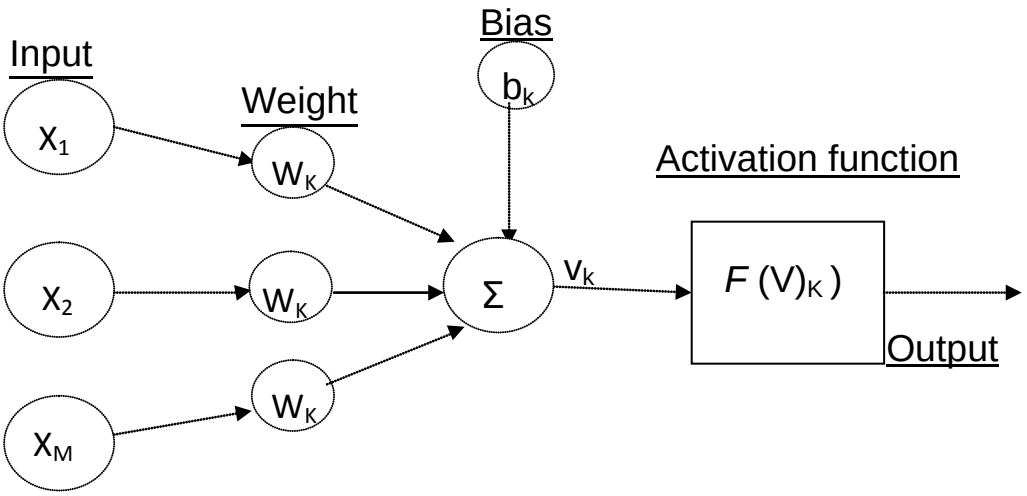


Fig. 2 : Mathematical structure of ANN (Ibeh et al, 2012)

V. Results and Discussions

Table 2 is obtained from the SPSS output for the analysis of the multiple linear regressions relating the measured values of global solar radiation as a function of the maximum temperature (Tm) and the sunshine hour (n). The standard error for each of the variables is indicated in the table 3. From figure 1, the highest value of solar radiation is 18.16 and 18.11MJ/M^2/Day respectively. These results suggest that there is peak dry season in Warri during December and April when the solar radiation is high. Again, the low values of solar radiation occur from May to August, which indicates peak rainfall Warri when the sky is cloudy and solar radiation is low.

Table 2 : Standard error of the models

Variable	Coefficient Std.	Std. Error	*t*-statistic	Prob.
Tm	-0.362	0.656	-0.461	0.785
$\bar{n}$	0.966	0.428	0.830	0.617
Constant	22.591	0.301	1.098	0.000

ERROR	ANN	REGRESSION	ANGSTROM-PRESCOTT
MBE	0.0085	0.2490	0.0352
RMSE	0.0004	0.0104	0.0251
MPE	-1.2580	-2.2116	-1.6230

Table 3 : Summary of error of the two models

VI. Conclusion

In this paper, three techniques for modeling and forecasting the solar radiation of Warri Nigeria: Neural Network, Angstrom-Prescott and Statistical Technique. The forecasting ability of these models is accessed on the basis of MBE, RMSE and MPE. We have discovered the fact that Neural Networks output perform better in forecasting from table 2 and 3 compare to other two models. Thus, ANN should be used for prediction of global solar radiation of the location and other location that has similar condition.

References Références Referencias

1. Beenstock M, Goldin E, Nabot D (1990). Energy Economics. 21: 168-183.
2. Hair A and Tatham B (1998). Multivariate Data Analysis. Pearson Education press.
3. Massie DD (2001). Neural Network fundamentals for scientist and engineers, Proceedings of the international conference of efficiency, Costs, Optimization, Simulation and Environmental Aspects of Energy systems and Processes, Istanbul Turkey. (BCOS '01), pp. 123-128
4. Ibeh GF and Agbo GA (2012). Estimation of monthly global solar radiation for Warri-Nigeria, using Angstrom and ANN models. Advances in Science research, 3(1): 12-13.
5. Ibeh, GF, Agbo,GA, Rabia, S, and Chikwenze, AR (2012). Comparison of empirical and artificial neural network models for the correlation of monthly average global solar radiation with sunshine hours in Minna, Nigeria. International Journal of physical Sciences. Vol. 7(8), pp. 1162-1165
6. Igbal, M (1983). An Introduction to Solar Radiation. Academy Press: New York, NY
7. Klein, SA (1977). "Calculation of Monthly Average Isolation on Tilted Surfaces". *Solar Energy*. 19: 307-311
8. Mendenhall and Beaver (1994). Introduction to Probability and Statistics. Ninth Edition, International, Thomson Publishing.

9. NeuralWare, NeuralWorks Predict, Available: http://www.neuralware.com
10. Olaniyi SM, Anthony OA, Adeola OM, Lawrence M I, Chukwuemeka OA, Comparison of Regression Model and Artificial Neural Network Model for the prediction of Electrical Power generated in Nigeria, Advances in Applied Science Research, 2011, 2 (5):329-339
11. Sharda R, and Patil R (1990). Neural Networks as forecasting experts: an empirical test," Proceedings of the 1990 International Joint Conference on Neural Networks, Vol-I, pp. 491-494, Washington DC, USA.
12. SPSS17 software (2008) SPSS Inc., *http://www.spss.com*.

Global Journal of Computer Science and Technology (D) Volume XII Issue XI Version I Year 2012
12

This page is intentionally left blank

Global Journal of Computer Science and Technology
Neural & Artificial Intelligence
Volume 12 Issue 11 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

Mobile Robot for Object Detection Using Image Processing

By Himanshu Borse, Amol Dumbare, Rohit Gaikwad & Nikhil Lende
Pune University

Abstract - This paper describes a robotic application that tracks a moving object by utilizing a mobile robot with sensors and image processing. The robotic platform uses a visual camera to sense the movement of the desired object and a range sensor to help the robot detect and then avoid obstacles in real time while continuing to detect and follow the desired object. In terms of real-time obstacle avoidance capacity, this paper also presents an algorithm for this robotic application specifically. Experimental results show that the robotic and intelligent system can fulfill the requirements for detecting an object and avoiding obstacles simultaneously.

GJCST-E Classification : I.5.m

Strictly as per the compliance and regulations of:

Mobile Robot for Object Detection Using Image Processing

Himanshu Borse [α], Amol Dumbare [σ], Rohit Gaikwad [ρ] & Nikhil Lende [ω]

Abstract - This paper describes a robotic application that tracks a moving object by utilizing a mobile robot with sensors and image processing. The robotic platform uses a visual camera to sense the movement of the desired object and a range sensor to help the robot detect and then avoid obstacles in real time while continuing to detect and follow the desired object. In terms of real-time obstacle avoidance capacity, this paper also presents an algorithm for this robotic application specifically. Experimental results show that the robotic and intelligent system can fulfill the requirements for detecting an object and avoiding obstacles simultaneously.

I. Introduction

Video tracking, surveillance systems, and robotic platforms are fields that have been well studied in the past decade. However, in the majority of surveillance and video tracking systems, the sensors are stationary. The stationary systems require the desired object to stay within the surveillance range of the system. If the object goes beyond this range.

It no longer becomes tractable. One solution to this problem is to design the system as a mobile system that uses a infrared range sensor, and a visual-spectrum camera, to track the object and avoid obstacles . This research topic has been partially studied in several different areas. Studies made by the automotive industry in this area develop systems that assist a human driver for safety and comfort . NASA has applied this to help astronauts to carry more equipment while walking on the moon . These systems are primarily concerned with object tracking, and the obstacle avoidance problem.

The contributions of this paper are to present a mobile robotic system which can simultaneously detect an object and avoid obstacles in real-time. We first introduce the system architecture, then present strategy for object detection, obstacle detection, obstacle avoidance mechanism and robot control. Finally, the experiment and conclusion will be addressed.

II. System Architecture

In general, the overall system consists of five main phases: image input, object detection, obstacle detection, obstacle avoidance and robot mobility phase. If no obstacles are detected, then system skips the obstacle avoidance phase, and only uses four phases. The following sections explain how each phase works individually, and how the various phases work in conjunction with each other.

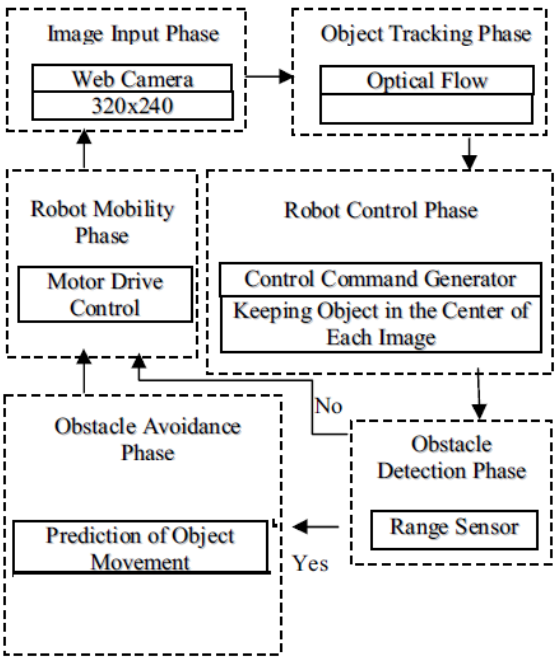


Figure 1 : This block diagram shows the overall system

a) Image Input Phase

The Logitech Web Camera has a fixed view and is attached to the robotic platform. It is used to acquire color 640x420 images. The camera is tasked to capture the object image.

b) Object Detection

In this system we have one webcam which is placed on the robot platform. Webcam captures the image and stores it on hard disk. In computer system preprocessing is done on the image to convert it into gray scale image. After this we find region for an object. On the basis of this region we compare the image with previously stored image. If the image matches with the one stored on the hard disk then robot moves in forward direction towards the object. If images doesn't match robot rotates using castroy wheel and repeats the process. First convert an image to gray scale, find gradient, boost the image. Then develop histogram from gray scaled image to find the region desired region of object.

E-mail α : hp2451@gmail.com
E-mail σ : amoldumbare12@gmail.com
E-mail ρ : rohitgai208@gmail.com
E-mail ω : nikhillende52@gmail.com

c) Robot Control Phase

If captured image matches with the image of an object then character "V" is sent indicating that required object has found & robot moves towards the object. If there is no match between captured image and destination image then character "A" is sent indicating that required object was not found. In such case robot takes a left turn. It keeps capturing and comparing the images until the captured image matches with the destination image.

d) Obstacle Detection Phase

IR sensor is mounted on the robot for obstacle detection . Whenever an IR sensor detects an obstacle, LED on IR sensor glows. Corresponding data is sent to microcontroller due to this robot moves backwards and then takes a left turn so that it can avoid the obstacle and find a new path.

e) Obstacle Avoidance Phase

Whenever an IR sensor detects an obstacle, LED on IR sensor glows. Corresponding data is sent to microcontroller due to this robot moves backwards and then takes a left turn so that it can avoid the obstacle and find a new path.

III. Platform

Figure shows the entire system including the web camera, IR sensor and robotic platform. Robot mobility is achieved through two wheels at back and one castroy wheel at front. Motion can be controlled directly by a computer system sending motion commands to the motors via RF module signal. Web Camera IR Sensor RF Module Computer.

a) Board

PIC16F877A microcontroller is used for robot computing system. Interfacing of microcontroller and LCD is done for validation purpose. The code to control is developed using MPLAB C. This software is retargeted specifically for PIC16F877A microcontroller.

b) Wireless Transmission

Wireless transmission is done through RF module CC2500 which is interfaced with microcontroller and connected to computer system.

c) Motor Driver System

Motor driver system consists of two DC motors and L239D motor driver IC.

d) Power Supply

Rechargeble batteries are used for power supply. Batteries provide clean reliable power supply and can be recharged. 12V power supply is required for robot to work and 9V is required for RF module to work. So we are using 12V and 9V batteries.

IV. Conclusion

Thus we have implemented Robotic Application which will detect the objects and avoid the obstacles. The application we have developed is a Desktop Application in that the user gives a command to capture image .This image is stored as destination image. After that user gives command to start the robot which will capture the image. This newly captured image will be compared with the destination image.Comparison will give a conclusion wheather the captured image matches with the destination image or not. This system compares the images using regionwise comparison. The system is implemented using J2EE, Java Swing and proteus technologies.

References Références Referencias

1. I. Haritaoglu, D. Harwood, and L. S. Davis, "W4: Real- Time Surveillance of People and Their Activities", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 22, No. 8, pp.809-830, August 2000.
2. D. Schulz, W. Burgard, D. Fox, and A. B. Cremers, "Tracking Multiple Moving Objects with a Mobile Robot", IEEE Conference on Computer Vision and Pattern Recognition 2001, Vol. 1, pp. 1-371, 2001.
3. O. Khatib, "Real-Time Obstacle Avoidance for Manipulators and Mobile Robots", Int. Journal of Robotics Research, 5(1), pp.90-98, 1986

Global Journal of Computer Science and Technology
Neural & Artificial Intelligence
Volume 12 Issue 11 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

Review on Financial Forecasting using Neural Network and Data Mining Technique

By Mehzabin Shaikh & Mrs. Gyankamal J. Chhajed
Faculty of Information Technology Polytechnic University of Tirana Tirana, Albania

Abstract - The rise of economic globalization and evolution of information technology, financial data are being generated and accumulated at an extraordinary speed. As a result, there has been a critical need for automated approaches to effective and efficient utilization of massive amount of financial data to support companies and individuals in strategic planning and investment decision-making. The competitive advantages achieved by data mining include increased revenue, reduced cost, and much improved marketplace responsiveness and awareness. There has been a large body of research and practice focusing on exploring data mining techniques to solve financial problems. This paper describes data mining in the context of financial application from both technical and application perspective by comparing different data mining techniques.

Keywords : *Data mining in Financial Forecasting, Data Mining Algorithm, Technical Indicators, BP neural network.*

GJCST-E Classification: *F.1.1*

Strictly as per the compliance and regulations of:

Review on Financial Forecasting using Neural Network and Data Mining Technique

Mehzabin Shaikh [α] & Mrs. Gyankamal J. Chhajed [σ]

Abstract - The rise of economic globalization and evolution of information technology, financial data are being generated and accumulated at an extraordinary speed. As a result, there has been a critical need for automated approaches to effective and efficient utilization of massive amount of financial data to support companies and individuals in strategic planning and investment decision-making. The competitive advantages achieved by data mining include increased revenue, reduced cost, and much improved marketplace responsiveness and awareness. There has been a large body of research and practice focusing on exploring data mining techniques to solve financial problems. This paper describes data mining in the context of financial application from both technical and application perspective by comparing different data mining techniques.

Keywords : Data mining in Financial Forecasting, Data Mining Algorithm, Technical Indicators, BP neural network.

I. Introduction

Data mining techniques have been used to uncover hidden patterns and predict future trends and behaviours in financial markets. The competitive advantages achieved by data mining include increased revenue, reduced cost, and much improved marketplace responsiveness and awareness. Data mining has been applied to a number of financial applications, including development of trading models, investment selection, loan assessment, portfolio optimization, fraud detection, bankruptcy prediction, real-estate assessment, and so on.

II. Literature Survey

a) *Classification and Issues of Data Mining in Financial Application*

Data mining aims to discover hidden knowledge, unknown patterns, and new rules from large databases that are potentially useful and ultimately understandable for making crucial decisions.

Based on the type of knowledge that is mined, data mining can be mainly classified into the following categories:

1) ***Association rule mining*** uncovers interesting correlation patterns among a large set of data items by showing attribute- value conditions that occur together frequently. A typical example is market basket analysis, which analyzes purchasing habits of customers by finding associations between different items in customers' —shopping baskets.‖

2) ***Classification and prediction*** is the process of identifying a set of common features and models that describe and distinguish data classes or concepts. The models are used to predict the class of objects whose class label is unknown. A bank, for example, may classify a loan application as either a fraud or a potential business using models based on characteristics of the applicant. A large number of classification models have been developed for predicting future trends of stock market indices and foreign exchange rates.

3) ***Clustering analysis*** segments a large set of data into subsets or clusters. Each cluster is a collection of data objects that are similar to one another within the same cluster but dissimilar to objects in other clusters. In other words, objects are clustered based on the principle of maximizing the intra-class similarity while minimizing the inter-class similarity.

4) ***Sequential pattern and time-series mining*** example, clustering techniques can be used to identify stable dependencies for risk management and investment management. looks for patterns where one event (or value) leads to another later event (or value). One example is that after the inflation rate increases, the stock market is likely to go down.

First, data mining needs to take ultimate applications into account. For example, credit card fraud detection and stock market prediction may require different data mining techniques.

Second, data mining is dependent upon the features of data. For example, if the data are of time series, data mining techniques should reflect the features of time sequence.

Third, data mining should take advantage of domain models. In finance, there are many well-developed models that provide insight into attributes that are important for specific applications. Many applications combine data mining techniques with various finance and accounting models (e.g., capital asset pricing model and the Kareken–Wallace model) [1].

Author α : M.E. student Department of Computer Engg, VPCOE, Baramati. E-mail : mehzabin.shaikh@gmail.com
Author σ : Asst. Professor and Head of Computer Engg, VPCOE, Baramati. E-mail : gjchhajed @gmail.com

b) Existing Data Mining Techniques

i. *Overview of Data Mining Techniques*

1. Neural Networks : Artificial neural networks are computer models built to emulate the human pattern recognition function through a similar parallel processing structure of multiple inputs. A neural network consists of a set of fundamental processing elements (also called neurons) that are distributed in a few hierarchical layers. Most neural networks contain three types of layers: input, hidden, and output.

2. Genetic Algorithms : The basic idea of genetic algorithms is that given a problem, the genetic pool of a specific population potentially contains the solution, or a better solution. Based on genetic and evolutionary principles, the genetic algorithm repeatedly modifies a population of artificial structures through the application of initialization, selection, crossover, and mutation operators in order to obtain an evolved solution.

3. Statistical Inference : Statistics provides a solid theoretical foundation for the problem of data analysis. Through hypothesis validation and/or exploratory data analysis, statistical techniques give asymptotic results that can be used to describe the likelihood in large samples. The basic statistical exploratory methods include such techniques as examining distribution of variables, reviewing large correlation matrices for coefficients that meet certain thresholds, and examining multidimensional frequency tables.

4. Rule Induction : Rule induction models belong to the logical, pattern distillation based approaches of data mining. Based on data sets, these techniques produce a set of if-then rules to represent significant patterns and create prediction models. Such models are fully transparent and provide complete explanations of their predictions. One commonly used and well-known type of rule induction is the family of algorithms that produce decision trees.

5. Data Visualization*—"Seeing" the Data:* Data are difficult to interpret due to its overwhelming size and complexity. In order to achieve effective data mining, it is important to include people in the data exploration process and combine the flexibility, creativity, and general knowledge of people with the enormous storage capacity and computational power of today's computers. Data visualization is the process of analyzing and converting data into graphics, thus taking advantage of human visual systems. large number of variables while still presenting useful information.

c) Applications of Data Mining In Finance

i. *Prediction of the Stock Market*

Investors in the market want to maximize their returns by buying or selling their investments at an appropriate time. Since stock market data are highly time-variant and are normally in a nonlinear pattern, predicting the future trend (i.e., rise, decrease, or remain steady) of a stock is a challenging problem. The dominant data mining technique used in stock market prediction so far is neural network modeling, including back-propagation (BP) networks, probabilistic neural networks, and recurrent neural networks The basic assumption is that similar input time series should produce similar output time series while ignoring intra-day fluctuations compared to regression models with a back-propagation network using the same data for stock prediction. Results showed that back-propagation network was a better predictor.

ii. *Portfolio Management*

Portfolio management is a major issue in investment. It concerns how individuals decide which securities to hold in investment portfolios and how funds should be allocated among broader asset classes, such as stocks versus bonds, and domestic securities versus foreign securities. The primary goal is to choose a set of risk assets to create a portfolio in order to maximize the return under certain risk or to minimize the risk for obtaining a specific return.

iii. *Bankruptcy Prediction*

Predicting bankruptcy is of great benefit to those who have some relations to a firm concerned, for bankruptcy is a final state of corporate failure. In the 21st Century, corporate bankruptcy in the world has reached an unprecedented level. It results in huge economic losses to companies, stockholders, employees, and customers, together with tremendous social and economical cost to the nation.

iv. *Foreign Exchange Market*

"Foreign Exchange" is the simultaneous buying of one currency and selling of another. The foreign exchange market is the largest financial market in the world, with a daily average turnover of over US$1 trillion. Data mining has been applied to identifying such technical trading rules. The neural networks were trained using more than 21 years of data to predict 1-day future spot rates for several nominal exchange rates, and achieved 58% accuracy for trading the British Pound and 57% accuracy for trading the German Mark.

v. *Fraud Detection*

Credit card transactions continue to grow, taking an ever larger share of the U.S. payment system and leading to a higher rate of stolen account numbers and subsequent losses by banks. According to Meridian Research, financial institutions lost more than US$1 billion in credit and debit card fraud in 2001. Therefore, fraud detection is becoming a central application area of data mining, which aims at searching for patterns indicative of fraud. Improving fraud detection is essential

to reducing the loss and maintaining the viability of the payment system.

vi. *Data Mining in Other Financial Applications*

In addition to the above applications that are discussed, above data mining techniques have also been applied to other financial applications such as loan risk analysis and payment prediction, mortgage scoring and real estate services. Data mining systems can determine whether or not a customer will be able to pay off their loans based on his/her income, age, and historical credit information, etc. Neural networks have been used for providing recommendations to grant or deny a loan based on financial ratios, past credit ratings, and loan records.

III. BP Neural Network

Artificial neural network is a large broad network with a number of processing units (neurons) connected. It is an abstract, simplified and simulation to human brain, and reflects the basic characteristics of the human brain. Generally, the neural network is the multi-layered network topology, including the input layer, hidden layer and output layer. In dozens of neural network models that were put forward, researchers often use the Hopfield network, BP network [2].

Hopfield network is the most typical feedback network model, it is one of the models which are most commonly studied now. The Hopfield network is the monolayer constituted by the same neuron, and is also a symmetrically connected associative network without learning function. It can implement the restriction optimization and associative memory BP network is the back-propagation network. It is a multi-layer forward network, learning by minimum mean square error. It is one of the most widely used networks. It can be used in the field of language integration, identification and adaptive control, etc.

BP network is semi supervised learning. First of all, artificial neural network needs to learn a certain learning criteria, and then it can work. Guidelines for e-learning (Electronic Learning) can be listed as below. If the result yielded by network is wrong, then the network should reduce the possibility of making the same mistake next time through learning.

Back propagation network with the following indicator can be used to develop the system which can be used for financial forecasting.

a) Economic indicator used as input to the BP neural network

i. *Moving average (MA)*

Moving average line is a statistical mean, which sums up stock price of certain days and gives out an average, and then connects them into a line to observe the price trend. The function of moving average is obtaining the average cost during a certain period, and we can use the average cost curve and the movement of daily closing price changes in the line analysis of the change of bull and bear to study and to determine possible changes in stock.

ii. *Random indicator (KDJ)*

There are a total of three lines standing for random indicators in the stock, namely K line, D line and the J line. Random indicator not only considers the highest price, the lowest price in the calculation period, but also takes into account of the random amplitude in the course of the fluctuation of stock price. Therefore, researchers always think that random indicator can more truly reflect the volatility of stock price, and it plays an important role in prompting.

$$K=2Kt\text{-}1 / 3+RSV/3$$

$$D=2\ Dt\text{-}1 /3+ K/3$$

$$RSV=100(Cn\text{-}Ln)(Hn\text{-}Ln)$$

In the formula: C represents the daily closing price means the lowest price that day; H stands for the highest price that day.

$$J = 3K - 2D$$

iii. *Moving Average Convergence/Divergence (MACD)*

The principle of MACD is to use the functions of the signs for aggregation and separation of fast moving average and slow moving average, in addition to double smoothing operation in order to study and determine the timing of buy and sell.

iv. *Relative Strength Index (RSI)*

Relative Strength Index is an indicator comparing the average of closing high and the average of closing low .We can use it to analyze the market and strength in order to forecast the future of the market trend.

RSI = [average of increase / (average of increase + average of decline)] *100

v. *On Balance Volume (OBV)*

OBV is the degree of the active investors in the stock market. If there are a lot of buyers and sellers, the stock prices and the volume of trade will rise, the atmosphere of the stock market is warm, commonly known as the bull market; if there are a few of buyers and sellers, the stock prices and the volume of trade will decline. It can be seen that the impact of the rise and fall of stock prices and the volume of trade is the OBV of stock. The volume of shares and stock price can also reflect the degree of the rise and fall of popularity of stock.

vi. *BIAS*

BIAS is the ratio between the application index and the moving average. Base on BIAS, it is possible to observe the degree which the stock price deviate from the moving average price to decide to buy or sell.

RSI = [average of increase / (average of increase +average of decline)] *100

vii. *Increase scope*

Increase scope= (the stock market closing price of today the stock market opening price of today) / the stock market opening price of today.

In order to avoid the situation that the dimensions of larger data have larger influence in the results than that of smaller data normalized the original data.

The formula of Normalization is as follows:

$$A\ ij\text{-}Min\ (Aj) / Max\ (Aj)\text{-}Min\ (Aj)$$

The data were normalized to between 0 and 1.

IV. Analysis

Even though different exiting techniques can be used for financial forecasting, but BP neural network will be the best to develop a software tool that can be used for financial forecasting.

References Références Referencias

1. Stock data analysis based on BP neural network Zhou Yixin (software engineering department of international college of Qingdao University) QingDao, ChinaJie Zhang.Information Engineering College of Qingdao.2010
2. Discovering Golden Nuggets: Data Mining in Financial Application by Dongsong Zhang and Lina Zhou-IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART C: APPLICATIONS AND REVIEWS, VOL. 34, NO. 4
    1. Fengjing Shao, Zhong qing yu. Principle and algorithm of data mining. Sinohydro Press, 2008
    2. Huadong Qin. Based on neural network forecasting stock market trends. Chinese dissertation database, 2009
    3. Zhijun Peng. Some new data mining method and its application in Chinese securities market. Chinese dissertation database, 2009

Global Journal of Computer Science and Technology
Neural & Artificial Intelligence
Volume 12 Issue 11 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

Intelligent Intrusion Detection in Computer Networks Using Fuzzy Systems

By Amin Einipour
Islamic Azad University, Andimeshk, Iran

Abstract - The Internet and computer networks are exposed to an increasing number of security threats. With new types of attacks appearing continually, developing flexible and adaptive security oriented approaches is a severe challenge. Intrusion detection is a significant focus of research in the security of computer systems and networks. The security of computer networks plays a strategic role in modern computer systems. In order to enforce high protection levels against threats, a number of software tools are currently developed.

In this paper, we have focused on intrusion detection in computer networks by combination of fuzzy systems and Particle Swarm Optimization (PSO) algorithm. Fuzzy rules are desirable because of their interpretability by human experts. PSO algorithm is employed as meta-heuristic algorithm to optimize the obtained set of fuzzy rules. Results on intrusion detection dataset from KDD-Cup99 show that the proposed approach would be capable of classifying instances with high accuracy rate in addition to adequate interpretability of extracted rules.

Keywords : *Intrusion Detection, Fuzzy rule extraction, Particle Swarm Optimization (PSO) algorithm.*

GJCST-E Classification: *C.2.0*

Strictly as per the compliance and regulations of:

Intelligent Intrusion Detection in Computer Networks Using Fuzzy Systems

Amin Einipour

Abstract - The Internet and computer networks are exposed to an increasing number of security threats. With new types of attacks appearing continually, developing flexible and adaptive security oriented approaches is a severe challenge. Intrusion detection is a significant focus of research in the security of computer systems and networks. The security of computer networks plays a strategic role in modern computer systems. In order to enforce high protection levels against threats, a number of software tools are currently developed.

In this paper, we have focused on intrusion detection in computer networks by combination of fuzzy systems and Particle Swarm Optimization (PSO) algorithm. Fuzzy rules are desirable because of their interpretability by human experts. PSO algorithm is employed as meta-heuristic algorithm to optimize the obtained set of fuzzy rules. Results on intrusion detection dataset from KDD-Cup99 show that the proposed approach would be capable of classifying instances with high accuracy rate in addition to adequate interpretability of extracted rules.

Keywords : Intrusion Detection, Fuzzy rule extraction, Particle Swarm Optimization (PSO) algorithm.

I. Introduction

Data mining usually means the methodologies and tools for the efficient new knowledge discovery from databases. It is also a form of knowledge discovery essential for solving problems in a specific domain.

An intrusion is defined as any set of actions that attempt to compromise the integrity, confidentiality, or availability of a resource [1]. An Intrusion Detection System (IDS) monitors and restricts user access to the computer system by applying certain rules. These rules are based on expert knowledge extracted from skilled administrators, who construct attack scenarios and apply them to find system exploits. The system identifies all intrusions by users and takes or recommends necessary action to stop an attack on the database.

Two approaches to intrusion detection are currently used. The first one, called misuse detection, is based on attack signatures, i.e., on a detailed description of the sequence of actions performed by the attacker. This approach allows the detection of intrusions matching perfectly the signatures, so that new attacks performed by slight modification of known attacks cannot be detected. The second approach is based on statistical knowledge about the normal activity of the computer system, i.e., a statistical profile of what constitutes the legitimate traffic in the network. In this case, intrusions correspond to anomalous network activity, i.e. to traffic whose statistical profile deviates significantly from the normal one [2, 3].

Earlier studies, the statistical related techniques were most commonly used data mining approaches to construct classification models. However, as the intrusion detection classification problem is highly nonlinear in nature, it is hard to develop a comprehensive model taking into account all the independent variables using conventional statistical modeling techniques. Furthermore, traditional ad hoc mixtures of statistical techniques and data management tools are no longer adequate for analyzing the vast collection of data. For the needs of improving the prediction accuracy in intrusion detection, more and more researchers have tried to apply artificial intelligence related approaches for intrusion detection in computer networks [4, 5].

A good computerized detection tool should possess two characteristics. First, the tool must attain the highest possible performance. Moreover, it would be highly desirable to be in possession of a so-called degree of confidence: the system not only provides a crisp detection, but also outputs a numeric value that represents the degree to which the system is confident about its response. Second, it would be highly beneficial for such a detection system to be human-friendly, exhibiting so-called interpretability. This means that the experts in computer networks is not faced with a black box that simply spouts answers (albeit correct) with no explanation; rather, we would like for the system to provide some insight as to how it derives its outputs.

Some experimental studies reported that success of artificial neural networks in intrusion detection [6-8], but there is a major drawback in building and using a model in which the user cannot readily comprehend the final rules that neural networks models acquire. In other words, the results of training a neural network are internal weights distributed throughout the network. These weights provide no more insight into why the solution is valid than asking many human experts why a particular decision is the right decision. For example, the weights are not readily understandable although, increasingly, sophisticated techniques for probing into neural networks help provide some explanation. It is also some recently studies used

Author : Department of Computer, Andimeshk Branch, Islamic Azad University, Andimeshk, Iran. E-mail : a.einipour@gmail.com

meta-heuristic and K-NN approaches to intrusion detection which interpretability of these approaches higher than of neural networks [9-12, 21].

In this paper we combine two methodologies—fuzzy systems and meta-heuristic algorithm—so as to automatically produce systems for intrusion detection in computer networks. The major advantage of fuzzy systems is that they favor interpretability; however, finding good fuzzy systems can be quite an arduous task. This is where PSO algorithm step in, enabling the automatic production of fuzzy systems, based on a database of training cases. Our fuzzy-PSO approach produces systems exhibiting two prime characteristics: first, they attain high classification performance, with the possibility of attributing a confidence measure to the output detection; second, the resulting systems involve a few simple rules, and are therefore (human-) interpretable.

We believe the development of PSO algorithm for data mining is a promising research area, due to the following rationale.

This paper is organized as follows. In Section 2 we describe the Intrusion detection problem and KDD99 dataset, which is the focus of our interest in this paper. The third section describes the fuzzy systems based classification. Section 4 describes our particular PSO-Fuzzy approach to the intrusion detection problem. The fifth section reports on computational results evaluating the performance of the proposed system. Finally, the sixth section concludes the paper.

II. The Intrusion Detection Problem and KDD Cup 99 Dataset

The first major work in the area of intrusion detection was discussed by J.P Anderson in [13]. Anderson introduced the concept that certain types of threats to the security of computer systems could be identified through a review of information contained in the system‘s audit trail. Many types of operating systems, particularly the various ―flavors‖ of UNIX, automatically create a report which details the activity occurring on the system. Anderson identified three threats which could be identified from a concentrated review of the audit data:

1. External Penetrations - Unauthorized users of the system.
2. Internal Penetrations - Authorized system users who utilize the system in an unauthorized manner.
3. Misfeasors - Authorized user who mislead their access privileges.

Anderson indicated that a particular class of external attackers, known as clandestine users, were particularly dangerous to the system resources. Clandestine users are those who evade both system access controls and auditing mechanisms through the manipulation of system privileges or by operating at a level that is lower than what is regularly monitored by the audit trail. Anderson suggested that clandestine users could be detected by lowering the level which is monitored by the audit trail, monitoring the functions that turn off the audit systems, or through a comparison of defined ―normal‖ usage patterns of system resource usage with those levels which are currently observed.

While the concept of manually reviewing operating system audit records for indications of intrusions was recognized as an extremely inefficient method of securing a computer system, Anderson‘s article served to initiate research into the area of intrusion detection. Subsequent research involved the development of automated techniques for the review of audit record data. Until recently, most intrusion detection mechanisms were based on an automated approach to Anderson's concepts. However, the recent development of new intrusion detection approaches and, more significantly, the necessary application of intrusion detection technologies to networked environments, is changing the focus of intrusion detection research.

Dr. Dorothy Denning proposed an intrusion detection model in 1987 which became a landmark in the research in this area [14]. The model which she proposed forms the fundamental core of most intrusion detection methodologies in use today. Because of the applicability of these concepts to most accepted intrusion detection systems, an overview of the primary concepts of the model are presented here to provide a basis of understanding the core technology.

Any statistical intrusion detection methodology requires the use of a set of definable metrics. These indices are the elements upon which all of the tool's statistical analysis is based. These metrics characterize the utilization of a variety of system resources. The resources which would be used in the definition of the metrics are required to be system characteristics which can be statistically based, (i.e., CPU usage, number of files accessed, number of login attempts).

These metrics are usually one of three different types. Event counters identify the occurrences of a specific action over a period of time. These metrics may include the number of login attempts, the number of times that a file has been accessed, or a measure of the number of incorrect passwords that are entered.

The second metric, time intervals, identify the time interval between two related events. Each time interval compares the delay in occurrence of the same or similar event An example of a time interval metric is the periods of time between a user‘s logins.

Finally, resource measurement is the concept of quantifying the amount of resources used by the system over a given period of time. Resource measurement incorporates individual event counters and time interval metrics to quantify the system. Examples of resource measurements include the expenditure of CPU time,

number of records written to a database, or the number of files transmitted over the network.

While not normally considered with the "traditional" intrusion detection metrics, keystroke dynamics is another method of quantifying a user's activities which offers an effective measure of user identification. The concept involves the development of an electronic signature of a user based on their individual typing characteristics. These characteristics usually include typing speed, intervals in typing, number of errors, and the user's typing rhythm. These characteristics may be verified on login and/or monitored throughout a session. Complete intrusion detection mechanisms have been developed exclusively around the use of keystroke dynamics techniques. [15]

In 1998, DARPA in concert with Lincoln Laboratory at MIT launched the DARPA 1998 dataset for evaluating IDS [16]. The DARPA 1998 dataset contains seven weeks of training and also two weeks of testing data. In total, there are 38 attacks in training data as well as in testing data. The refined version of DARPA dataset which contains only network data (i.e. Tcpdump data) is termed as KDD dataset [17]. The Third International Knowledge Discovery and Data Mining Tools Competition were held in colligation with KDD-99, the Fifth International Conference on Knowledge Discovery and Data Mining. KDD dataset is a dataset employed for this Third International Knowledge Discovery and Data Mining Tools Competition. KDD training dataset consists of relatively 4,900,000 single connection vectors where each single connection vectors consists of 41 features and is marked as either normal or an attack, with exactly one particular attack type [18]. These features had all forms of continuous and symbolic with extensively varying ranges falling in four categories:

- In a connection, the first category consists of the intrinsic features which comprises of the fundamental features of each individual TCP connections. Some of the features for each individual TCP connections are duration of the connection, the type of the protocol (TCP, UDP, etc.) and network service (http, telnet, etc.).
- The content features suggested by domain knowledge are used to assess the payload of the original TCP packets, such as the number of failed login attempts.
- Within a connection, the same host features observe the recognized connections that have the same destination host as present connection in past two seconds and the statistics related to the protocol behavior, service, etc are estimated.
- The similar same service features scrutinize the connections that have the same service as the current connection in past two seconds.

A variety of attacks incorporated in the dataset fall into following four major categories:

Denial of Service Attacks (DOS): A denial of service attack is an attack where the attacker constructs some computing or memory resource fully occupied or unavailable to manage legitimate requirements, or reject legitimate users right to use a machine.

User to Root Attacks(U2R): User to Root exploits are a category of exploits where the attacker initiate by accessing a normal user account on the system (possibly achieved by tracking down the passwords, a dictionary attack, or social engineering) and take advantage of some susceptibility to achieve root access to the system.

Remote to User Attacks (R2L): A Remote to User attack takes place when an attacker who has the capability to send packets to a machine over a network but does not have an account on that machine, makes use of some vulnerability to achieve local access as a user of that machine.

Probes (PRB): Probing is a category of attacks where an attacker examines a network to collect information or discover well-known vulnerabilities. These network investigations are reasonably valuable for an attacker who is staging an attack in future. An attacker who has a record, of which machines and services are accessible on a given network, can make use of this information to look for fragile points.

Table1 illustrates a number of attacks falling into four major categories and table 2 presents a complete listing of a set of features characterized for the connection records.

Table 1 : Various types of attacks described in four major categories

Denial of Service Attacks(DOS)	Back, land, neptune, pod, smurf, teardrop
User to Root Attacks(U2R)	Buffer _overflow, loadmodule, perl, rootkit,
Remote to Local Attacks(R2L)	Ftp_write, guess_passwd, imap, multihop, phf, spy, warezclient, warezmaster
Probes(PRB)	Satan, ipsweep, nmap, portsweep

Table 2 : A complete list of features given in KDD cup 99 dataset

Feature index	feature name	description	Type
1	duration	length (number of seconds) of the connection	Continuous
2	Protocol_type	type of the protocol, e.g. tcp, udp, etc.	Symbolic
3	service	network service on the destination, e.g., http, telnet,etc.	Symbolic
4	flag	normal or error status of the connection	Symbolic
5	Src_Bytes	number of data bytes from destination to source	Continuous
6	Dst_Bytes	number of data bytes from destination to source	Continuous
7	Land	1 if connection is from/to the same host/port; 0 otherwise	Symbolic
8	wrong_fragment	number of ``wrong'' fragments	Continuous
9	Urgent	number of urgent packets	Continuous
10	hot	number of ``hot'' indicators	Continuous
11	num_failed_logins	number of failed login attempts	Continuous
12	logged_in	1 if successfully logged in; 0 otherwise	Symbolic
13	num_compromised	number of ``compromised'' conditions	Continuous
14	root_shell	1 if root shell is obtained; 0 otherwise	Continuous
15	su_attempted	1 if ``su root'' command attempted; 0 otherwise	Continuous
16	num_root	number of ``root'' accesses	Continuous
17	num_file_creations	number of file creation operations	Continuous
18	num_shells	number of shell prompts	Continuous
19	num_access_files	number of operations on access control files	Continuous
20	num_outbound_cmds	number of outbound commands in an ftp session	Continuous
21	is_hot_login	1 if the login belongs to the ``hot'' list; 0 otherwise	Symbolic
22	is_guest_login	1 if the login is a ``guest'' login; 0 otherwise	Symbolic
23	Count	number of connections to the same host as the current connection in the past two seconds	Continuous
24	Srv_count	number of connections to the same service as the current connection in the past two seconds	Continuous
25	serror_rate	% of connections that have ``SYN'' errors	Continuous
26	srv_serror_rate	% of connections that have ``SYN'' errors	Continuous
27	rerror_rate	% of connections that have ``REJ'' errors	Continuous
28	srv_rerror_rate	% of connections that have ``REJ'' errors	Continuous
29	same_srv_rate	% of connections to the same service	Continuous
30	diff_srv_rate	% of connections to different services	Continuous
31	srv_diff_host_rate	% of connections to different hosts	Continuous
32	dst_host_count	count for destination host	Continuous
33	dst_host_srv_count	srv_count for destination host	Continuous
34	dst_host_same_srv_rate	same_srv_rate for destination host	Continuous
35	dst_host_diff_srv_rate	diff_srv_rate for destination host	Continuous
36	dst_host_same_src_port_rate	same_src_port_rate for destination host	Continuous
37	dst_host_srv_diff_host_rate	diff_host_rate for destination host	Continuous
38	dst_host_serror_rate	serror_rate for destination host	Continuous
39	dst_host_srv_serror_rate	srv_serror_rate for destination host	Continuous
40	dst_host_rerror_rate	rerror_rate for destination host	Continuous
41	dst_host_srv_rerror_rate	srv_serror_rate for destination host	Continuous

III. Fuzzy System Based Classification

Let us assume that our pattern classification problem is a c -class problem in the $n-$dimensional pattern space with continuous attributes. We also assume that m real vectors $x_p = (x_{p1}, x_{p2}, ..., x_{pn}), p = 1, 2, ..., m,$ are given as training patterns from the c classes ($c \ll m$).

Because the pattern space is $[0,1]^n$, attribute values of each pattern are $x_{pi} \in [0,1]$ for $p = 1, 2, ..., m$ and $i = 1, 2, ..., n$. In computer simulations of this paper, we normalize all attribute values of each data set into the unit interval [0, 1].

In the presented fuzzy classifier system, we use fuzzy if-then rules of the following form:

Rule R_j : If x_1 is A_{j1} and … and x_n is A_{jn} , then Class C_j with $CF = CF_j$.

where R_j is the label of the *jth* fuzzy if–then rule, $A_{j1}, ..., A_{jn}$ are antecedent fuzzy sets on the unit interval [0,1] , C_j is the consequent class (i.e., one of the given c classes), and CF_j is the grade of certainty of the fuzzy if–then rule R_j . In computer simulations, we use a typical set of linguistic values in Fig. 1 as antecedent fuzzy sets. The membership function of each linguistic value in Fig. 1 is specified by homogeneously partitioning the domain of each attribute into symmetric triangular fuzzy sets. We use such a simple specification in computer simulations to show the high performance of our fuzzy classifier system, even if the membership function of each antecedent fuzzy set is not tailored. However, we can use any tailored membership functions in our fuzzy classifier system for a particular pattern classification problem.

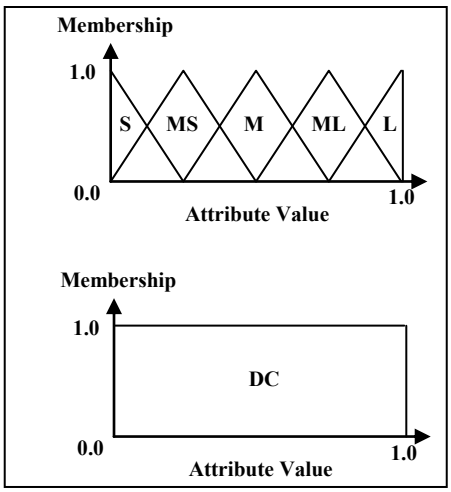


Fig. 1 : The used antecedent fuzzy sets in this paper. 1: Small, 2: medium small, 3: medium, 4: medium large, 5: large, and 6: don't care

The total number of fuzzy if–then rules is in the case of the -dimensional pattern classification problem. It is impossible to use all the fuzzy if– then rules in a single fuzzy rule base when the number of attributes (i.e.) is large (e.g., intrusion detection problem which n = 41). 6n n 6n n

Our fuzzy classifier system searches for a relatively small number of fuzzy if–then rules with high classification ability. Since the consequent class and the certainty grade of each fuzzy if–then rule can be determined from training patterns by a simple heuristic procedure [19], the task of our fuzzy classifier system is to generate combinations of antecedent fuzzy sets for a set of fuzzy if–then rules. While this task seems to be simple at first glance, in fact it is very difficult for high-dimensional pattern classification problems, since the search space involves combinations. 6n

In our fuzzy classifier system, the consequent Class and the grade of certainty of each fuzzy if– then rule are determined by a modified version of the heuristic procedure which is discussed in [18][19]. j C j CF

IV. Proposed Pso-Based Fuzzy System Approach

Swarm intelligence describes the collective behavior of decentralized, self organized natural or artificial systems. Swam intelligence model were employed in artificial intelligence. The expression was introduced in the year 1989 by Jing wang and Gerardo Beni in cellular robotic systems. Swarm Intelligence (SI) was a innovative pattern for solving optimizing problems. SI systems are typically made up of populations of simple agents interacting locally with one another and with their environment. The agent follows simple rules and the interactions between agents lead to the emergence of *"intelligent"* global behavior, unknown to the individual agents. Examples of SI include ant colonies, bird flocking, animal herding, bacterial growth and fish schooling.

The example algorithms of Swarm Intelligence are i) Ant Colony Optimization ii) Particle Swarm Optimization iii) Gravitational Search Algorithm iv) Stochastic diffusion search. Particle Swarm Optimization belongs to the class of swarm intelligence techniques that are used to resolve the optimization problems.

Particle Swarm Optimization (PSO) works with a population-based heuristic inspired by the social behavior of bird flocking aiming to find food [20]. In Particle Swarm Optimization the system initializes with a set of solutions and searches for optima by updating generations. The set of possible solutions is a set of particles, called swarm, which moves in the search space, in a cooperative search procedure. These moves are performed by an operator called velocity of a particle and moves it through an n-dimensional space based on the best positions of their leader (social component) and on their own best position (local component).

The main strength of PSO is its fast convergence, which compares with many global optimization algorithms like Genetic algorithms, Simulated Annealing and other global optimization algorithms. Particle Swarm Optimization shares many similarities with evolutionary computation techniques such as Genetic Algorithms. The system is initialized with a population of random solutions and searches for optima by updating generations. PSO has no evolution operators such as cross over and mutation. In PSO, the potential solutions called particles fly through the problem space by following the current optimum particles. PSO is a global optimization algorithm for dealing with problems in which a best solution can be represented as a point or surface search in n-dimensional space.

Hypotheses are plotted in this space and seeded with an initial velocity as well as a communication channel between the particles. Particles then move through the solution space and are evaluated according to some fitness criterion following each time step. The particles were accelerated in the direction of communication grouping which have better fitness values. The main advantage of such approach great global minimization strategies such as simulated annealing is that the large number of members that make up the particle swarm formulate the technique impressively flexible to the problem of local minima.

Each particle keeps track of its coordinates in the problem space which are associated with the best solution it has achieved so far. This value is called pbest. Another best value that is tracked by the particle swarm optimizer is the best value, obtained so far by any particle in the neighbors of the particle. This location is called lbest. When a particle takes all the population

as its topological neighbors, the best value is a global best and is called gbest.

The basic idea of combining PSO with data mining is simple. To extract this knowledge, a database may be considered as a large search space and a mining algorithm as a search strategy. PSO makes use of particles moving in an n-dimensional space to search for solutions for an n-variable function optimization problem. The datasets are the sample space to search and each attribute is a dimension for the PSO-miner.

The pseudo-code of the PSO-Based Search algorithm is presented in figure 2. Note that the input of this algorithm is a fuzzy rule and the output is an improved version of that fuzzy rule.

```
For each particle
Initialize particle
End For
Do
For each particle
Calculate fitness value of the particle fp
/*updating particle's best fitness value so far)*/
If fp is better than pBest
set current value as the new pBest
End For
/*updating population's best fitness value so far)*/
Set gBest to the best fitness value of all particles
For each particle
Calculate particle velocity according equation (1)
Update particle position according equation (2)
End For
While maximum iterations OR minimum error criteria
is not attained
```

Fig. 2 : Pseudo-code for the PSO-Based Search algorithm

a) Proposed Fuzzy-PSO Approach for Intrusion Detection in Computer Networks

Outline of the proposed approach for intrusion detection follows in figure 3.

A fuzzy inference system is a rule-based system that uses fuzzy logic, rather than Boolean logic, to reason about data [23]. Its basic structure includes four main components, as depicted in Fig. 3: (1) a fuzzifier, which translates crisp (real-valued) inputs into fuzzy values; (2) an inference engine that applies a fuzzy reasoning mechanism to obtain a fuzzy output; (3) a defuzzifier, which translates this latter output into a crisp value; and (4) a knowledge base, which contains both an ensemble of fuzzy rules, known as the rule base, and an ensemble of membership functions, known as the database.

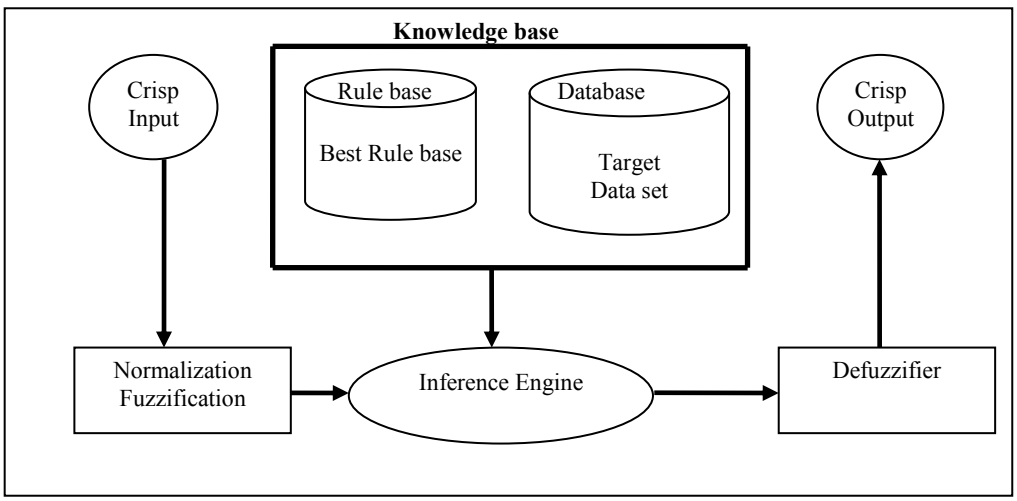


Fig. 3 : Basic structure of a fuzzy inference system for intrusion detection

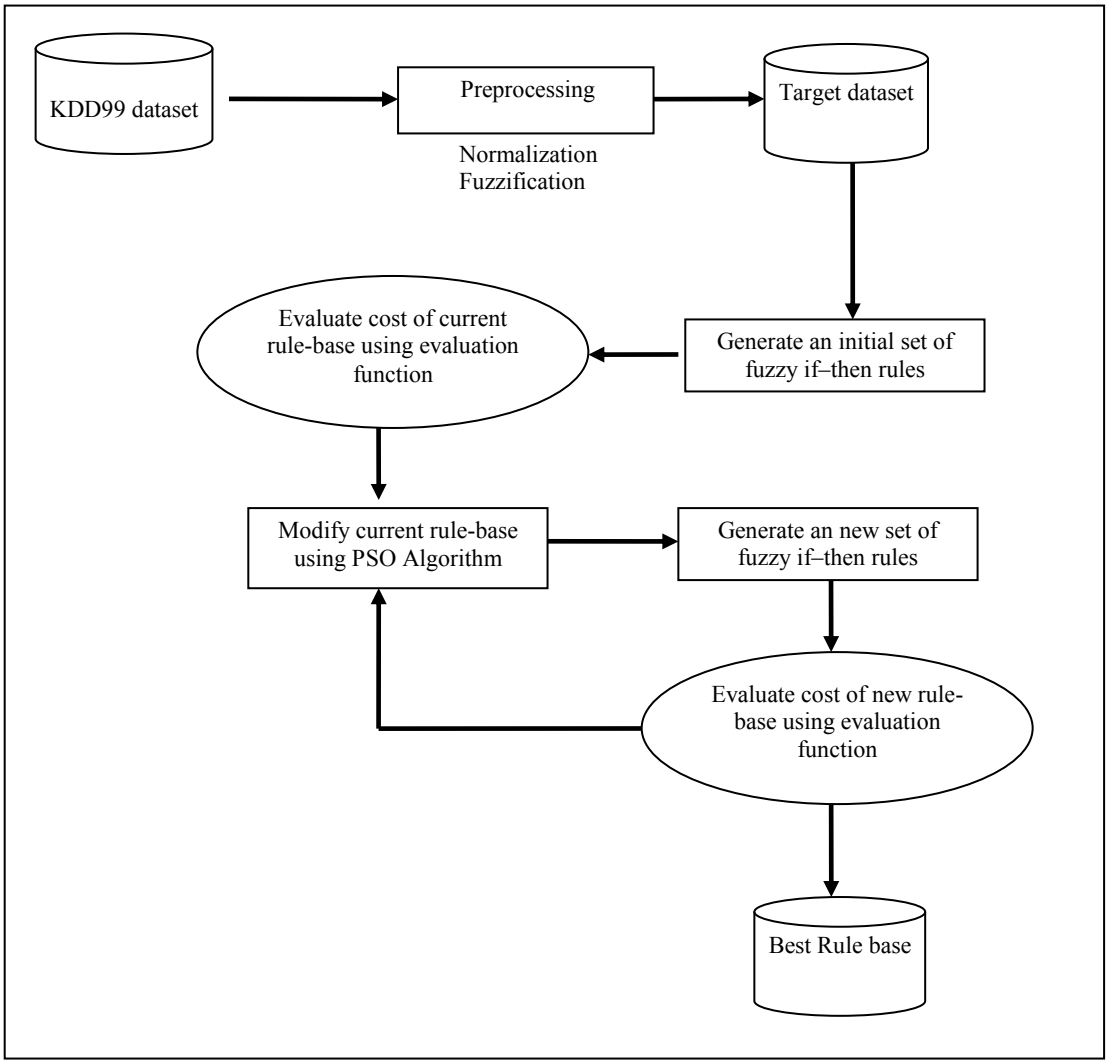


Fig. 4 : Generate Best Knowledge base Using PSO

The decision-making process is performed by the inference engine using the rules contained in the rule base. These fuzzy rules define the connection between input and output fuzzy variables.

Fuzzy modeling is the task of identifying the parameters of a fuzzy inference system so that a desired behavior is attained [24]. With the direct approach a fuzzy model is constructed using knowledge from a human expert. This task becomes difficult when the available knowledge is incompleteor when the problem space is very large, thus motivating the use of automatic approaches to fuzzy modeling.

There are several approaches to fuzzy modeling, based on neural networks [25, 26, 27], genetic algorithms [28, 29], and hybrid methods [9].

In this paper we use Particle Swarm Optimization (PSO) as automatic approach to produce knowledge base which indicated in figure 4.

In the next subsection we describe some steps of the proposed approach which presented in figure 3, 4.

i. *Generate initial rule-base*

There are different method for generate initial rule-base. One of the methods is that we assigned one of the symbols in figure 1 randomly. One method is that we increase probability of don't care fuzzy term. One innovative method is that we generate fuzzy if-then rules directly using training set. First, compatible fuzzy if-then rule is created; then antecedent part of some rules is replaced with don't care term.

The proposed evolutionary fuzzy system (EFS) is considered for each of the classes of the classification problem separately. One of the important benefits of this separation is that the learning system can focus on each of the classes of the classification problem. According to this fact, the mentioned random pattern is extracted according to the patterns of the training dataset, which their consequent class is the same as the class that the algorithm works on. Next, for this random pattern, we determine the most compatible combination of antecedent fuzzy sets using only the five linguistic

values (Figure. 1). The compatibility of antecedent fuzzy rules with the random pattern is calculated by equation (1).

$$\mu_j(x_p) = \mu_{j1}(x_{p1}) \times \ldots \times \mu_{jn}(x_{pn}) \qquad p = 1,2,\ldots,m \qquad (1)$$

Where $\mu_{A_{ji}}(.)$ is the membership function of Aji.

After generating each fuzzy if-then rule, the consequent class of this rule is determined according to (2).

$$\beta_{Class\ h}(R_j) = \sum_{x_p \in Class\ h} \mu_{R_j}(x_p) \qquad ,h = 1,2,\ldots,c \quad (2)$$

Where,

$$\beta_{Class\ C_j}(R_j) = \max\{\beta_{Class\ 1}(R_j),\ldots,\beta_{Class\ c}(R_j)\}. \quad (3)$$

Where β Class h (R j) is the sum of the compatibility grades of the training patterns in Class h with the fuzzy if–then rule Rj and N Class h is the number of training patterns which their corresponding class is Class h . Each of the fuzzy rules in the final classification has a certainty grade, which denotes the strength of that fuzzy rule. This number is calculated according to (4).

$$CF_j = \frac{\beta_{Class\ C_j}(R_j) - \overline{\beta}}{\sum_{h=1}^{c} \beta_{Class\ h}(R_j)} \qquad (4)$$

Where,

$$\overline{\beta} = \frac{\sum_{h \neq C_j} \beta_{Class\ h}(R_j)}{c-1} \qquad (5)$$

ii. *Evaluate rule-base*

The generation of each fuzzy rule is accepted only if its consequent class is the same as its corresponding random pattern class. Otherwise, the generated fuzzy rule is rejected and the rule generation process is repeated. After generation of N pop fuzzy if-then rules, the fitness value of each rule is evaluated by classifying all the given training patterns using the set of fuzzy if–then rules in the current population. The fitness value of the fuzzy if–then rule is evaluated by the following fitness function:

$$fitness(R_j) = NCP(R_j) \qquad (6)$$

Where $NCP(R_j)$ denotes the number of correctly classified training patterns by rule R_j .

iii. *PSO-based rule set update and optimization*

As we have mentioned in the previous subsections, initial rule-base is generated randomly, so that accuracy of the initial rule-base is low. For optimize initial rule-base, we use an PSO algorithm.

Input of this algorithm is a fuzzy rule and the output is an improved version of that fuzzy rule. The improvement is accomplished by some modifications (local search) to the current (input) fuzzy rule. The algorithm is capable of searching for the best modification according to the lifetime of the current fuzzy rule. in each step the algorithm performs one changes to the current (input) fuzzy rule. For each one value, a complete PSO process is done.

In PSO there are many fitness functions. By exploring Pareto dominance concepts, it is possible to obtain results with specific properties. Based on this concept each particle of the swarm could have different leaders, but only one may be selected to update the velocity. This set of leaders is stored in a repository, which contains the best non-dominated solutions found. The PSO components are defined as follows [22].

Each particle pi, at a time step t, has a position x (t) ϵ R^n, that represents a possible solution. The position of the particle, at time t +1, is obtained by adding its velocity, v (t) ϵ R^n, to x (t):

$$x(t+1) = x(t) + v(t+1) \qquad (9)$$

The velocity of a particle pi is based on the best position already fetched by the particle, pbest(t), and the best position already fetched by the set of neighbors of pi, Rh(t), that is a leader from the repository. The velocity update function, in time step t +1 is defined as follows:

$$v(t+1) = \varpi * v(t) + (c_1 * \phi_1) * (p_{best}(t) - x(t)) + (c_2 * \phi_2) * (R_h(t) - x(t)) \qquad (10)$$

The variables ϕ1and ϕ2, in Equation 10, are coefficients that determine the influence of the particle's positions. The constants c_1 and c_2 indicates how much each component influences on the velocity. The coefficient ϖ is the particle inertia and controls how much the previous velocity affects the current one R_h is a particle from the repository, chosen as a guide of p_i. There are many ways to make this choice. At the end of the algorithm, the solutions in the repository are the final output.

V. Experimental Result

Experiments were carried out on a subset of the database created by DARPA in the framework of the 1998 Intrusion Detection Evaluation Program [16]. We used the subset that was pre-processed by the Columbia University and distributed as part of the UCI KDD Archive [17]**Error! Reference source not found.** The available database is made up of a large number of network connections related to normal and malicious traffic. Each connection is represented with a 41-dimensional feature vector which presented in table 2. Connections are also labeled as belonging to one out of

five classes. One of these classes is the normal class and the rest indicates four different intrusion classes: PRB, DOS, U2R, and R2L which presented in table 1. These intrusion classes are a classification of 22 different types of attacks in a computer network.

This approach is implemented by using C++ programming language. We use 10-CV technique for evaluate proposed approach. In this technique, KDD99 dataset divided to 10 parties, nine parties for train set and one party for test set.

Table 3 is the confusion matrix of Proposed approach. The top-left entry of Table 3 shows that 3194 instances of the actual PRB test set were detected to be PRB; the last column indicates that 76.66% of the actual PRB samples were detected correctly. In the same way, for R2L, 1971 instances of the actual _attack' test set were correctly detected. The last column indicates that 12.17% of the actual R2L samples were detected correctly. The bottom row shows that 83.48% of the test set classified, as R2L indeed belongs to R2L. The bottom-right entry of the table 3 shows that 93.70% of all patterns in the test set are correctly classified.

Table 4 represents the cost matrix that defines the cost for each type of misclassification.

We aim at minimizing that cost function. Given the confusion and cost matrixes, we calculated the cost of our simulated annealing based fuzzy intrusion detection system as shown in table 5. The bottom-right entry of the table 5 shows that the classification cost of our algorithm is 0.1872.

Proposed approach is compared with some algorithms, such as C4.5, *k*-NN, Naïve Bayes, SVM, MLP. Result of comparison is indexed in table 6.

Table 3 : Confusion Matrix for the Proposed Approach

Real Class	Detected Class				
	PRB	DOS	U2R	R2L	%
PRB	3194	306	0	0	76.66
DOS	576	226388	10	0	98.49
U2R	116	0	37	11	16.22
R2L	77	3447	8	1971	12.17
%	76.66	98.33	46.83	83.48	**93.70**

Table 4 : Cost Matrix Used To Evaluate the Confusion of Proposed Approach

Real Class	Detected Class			
	PRB	DOS	U2R	R2L
PRB	0	2	2	2
DOS	1	0	2	2
U2R	2	2	0	2
R2L	2	2	2	0

Table 5 : Cost-Based Scoring of the Proposed Approach

Real Class	Detected Class				
	PRB	DOS	U2R	R2L	
PRB	0	612	0	0	
DOS	576	0	20	0	
U2R	232	0	0	22	
R2L	154	6894	16	0	
					0.1872

Table 6 : Recall, Precision, and F-measure For Different Classifier. The Best Values are Bold

Class	Algorithm	C4.5	5-NN Error! Reference source not found.	SVM Error! Reference source not found.	Winner Entry	Proposed Approach
PRB	Recall	81.88	81.61	**86.27**	83.30	76.66
	Precision	52.20	55. 6	77.72	64.81	**76.66**
	F-measure	63.76	66.05	81.77	72.90	**86.2**
DOS	Recall	96.99	97.00	97.65	97.10	**99.20**
	Precision	99.69	99.42	99.86	**99.88**	98.39
	F-measure	98.32	98.19	98.70	98.47	**98.79**
U2R	Recall	14.47	14.91	10.09	13.20	**16.23**
	Precision	9.35	5.47	53.49	**71.43**	32.46
	F-measure	11.36	8.00	16.97	**22.28**	21.64
R2L	Recall	1.45	6.90	3.55	8.40	**12.65**
	Precision	30.32	66.97	62.39	**98.84**	90.26
	F-measure	2.77	12.51	6.71	15.48	**22.19**

VI. Conclusions

In this paper, we focused on intrusion detection in computer networks by combination of fuzzy systems and PSO algorithm. The proposed method performs the classification task and extracts required knowledge using fuzzy rule based systems which consists of fuzzy if-then rules. Particle Swarm Optimization algorithm is employed to optimize the obtained set of fuzzy rules. The proposed system has two main features of data mining techniques which are high reliability and adequate interpretability, and is comparable with several well-known algorithms. Results on *intrusion detection* data set from KDD cup-99 repository show that the proposed approach would be capable of classifying intrusion instances with high accuracy rate in addition to adequate interpretability of extracted rules.

References Références Referencias

1. Heady, G. Luger, A. Maccabe, and M. Servilla, -*The architecture of network level intrusion detection system*‖, echnical Report, Department of Computer Science, University of New Mexico, 1990.
2. J. McHugh, A. Christie, and J. Allen, —Defending Yourself: The Role of Intrusion Detection Systems‖, IEEE Software, Sept./Oct. 2000, pp. 42-51
3. P.E. Proctor, The Practical Intrusion Detection Handbook, Prentice Hall, 2001
4. Xu L., A. Krzyzak and C.Y. Suen, "*Methods for combining multiple classifiers and their applications to handwriting recognition*", *IEEE Trans. Systems, Man and Cybernetics* 22, 1992, pp. 418-435.
5. Anderson D, Lunt TF, Javitz H, Tamaru A, Valdes A. "Detecting unusual program behavior using the statistical component of the next-generation intrusion detection expert system (NIDES)", Menlo Park, CA, USA: Computer Science Laboratory, SRI International; 1995. SRIO-CSL-95-06.
6. A.K. Ghosh and A. Schwartzbard, —*A Study in Using Neural Networks for Anomaly and Misuse Detection*", Proc. of the USENIX Security Symposium, August 23-26, 1999, Washington, USA.
7. J. Cannady, —*An adaptive neural network approach to intrusion detection and response*‖, PhD Thesis, School of Comp. and Inf. Sci., Nova Southeastern University, 2000.
8. J.M. Bonifacio et al. —*Neural Networks applied in intrusion detection systems*‖, Proc. of the IEEE World congress on Comp. Intell. (WCCI '98), 1998.
9. Mohammad saniee abadeh, Jafar Habibi, " *A Hybridization of Evolutionary Fuzzy Systems and Ant Colony Optimization for Intrusion Detection* ", INTERNATIONAL JOURNAL OF INFORMATION SECURITY, (2010), Vol. 2, No. 1, pp. 33-45.
10. Mohammad saniee abadeh, hamid mohammadi, Jafar Habibi, "*Design and analysis of genetic fuzzy systems for intrusion detection in computer networks* ", EXPERT SYSTEMS WITH APPLICATIONS, (2011), Vol. 38, No. 6, pp. 7067-7075.
11. M Saniee, Jafar Habibi, C Lucas, "*Intrusion Detection Using a Fuzzy Genetics-Based learning algorithm*", JOURNAL OF NETWORK AND COMPUTER APPLICATIONS, (2007), No. 0, pp. 414-428.

12. Mohammad Saniee Abadeh, Jafar Habibi, Zeynab Barzegar, Muna Sergi, "*A Parallel Genetic Local Search Algorithm for Intrusion Detection in Computer Networks*", ENGINEERING APPLICATIONS OF ARTIFICIAL INTELLIGENCE, (2007), Vol. 20, No. 20, pp. 1058-1069.
13. Anderson, J.P., "*Computer Security Threat Monitoring and Surveillance*" , Technical Report, J.P. Anderson Company, Fort Washington, Pennsylvania, April, 1980.
14. Denning, Dorothy, "An Intrusion-Detection Model" , IEEE Transactions on Software Engineering, February, 1987 , Vol. SE-13, No. 2.
15. Lunt, T.F., "*Automated Audit Trail Analysis and Intrusion Detection: A Survey*", Proceedings of the 11th National Computer Security Conference, 1990.
16. http://www.ll.mit.edu/mission/communications/ist/corpora/ideval/data/1998data.html
17. http://www.sigkdd.org/kddcup/index.php?section=1999&method=data
18. Mahbod Tavallaee, Ebrahim Bagheri, Wei Lu and Ali A. Ghorbani, "*A detailed analysis of the KDD CUP 99 data set*", in Proceedings of the Second IEEE international conference on Computational intelligence for security and defense applications, pp. 53-58, Ottawa, Ontario, Canada, 2009.
19. H. Ishibuchi, K. Nozaki, and H. Tanaka, *—Distributed representation of fuzzy rules and its application to pattern classification*‖ , Fuzzy Sets and Systems, 52(1), 1992, pp. 21-32.
20. Kennedy, J. and Eberhart, R. (1995). Particle swarm optimization. In IEEE International Conference on Neural Networks, pages 1942–1948. IEEE Press.
21. Amin Einipour *—A Fuzzy-ACO Method for Detect Breast Cancer*‖ Global Journal of Health Science October 2011.
22. Andre B. de Carvalho, Taylor Savegnago, Aurora Pozo, *"Swarm Intelligence for Rule Discovery in Data Mining"*, pp 314-319, In ICEIS 2010, Volume 2, AIDSS, Funchal, Madeira, Portugal, June 8 - 12, 2010
23. Yager RR, Zadeh LA. Fuzzy Sets, Neural Networks, and Soft Computing. New York: Van Nostrand Reinhold, 1994.
24. Yager RR, Filev DP. Essentials of Fuzzy Modeling and Control. Wiley, 1994.
25. Jang J-S R, Sun C-T. , "*Neuro-fuzzy modeling and control. Proceedings*", IEEE. 1995 83(3):378–406.
26. Lee K-M, Kwak D-H, Lee-Kwang H. , "*Fuzzy inference neural network for fuzzy model tuning*" , IEEE Trans Syst Man Cybern 1996;26(4):637–45.
27. Lin C-T, Lee CSG. , "Reinforcement structure: parameter learning for neural-network-based fuzzy logic control systems ", IEEE Trans Fuzzy Syst 1994;2(1):46–63.
28. Heider H, Drabe T., "*Fuzzy system design with a cascaded genetic algorithm*" , Proceedings of 1997 IEEE International Conference on Evolutionary Computation. IEEE and IEEE Neural Network Council and Evolutionary Programming Society, 1997:585–588.
29. Nawa NE, Hashiyama T, Furuhashi T, Uchikawa Y. , " *A study on fuzzy rules discovery using pseudo-bacterial genetic algorithm with adaptive operator*" , Proceedings of 1997 IEEE International Conference on Evolutionary Computation. IEEE and IEEE Neural Network Council and Evolutionary Programming Society, 1997.

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

S

T

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