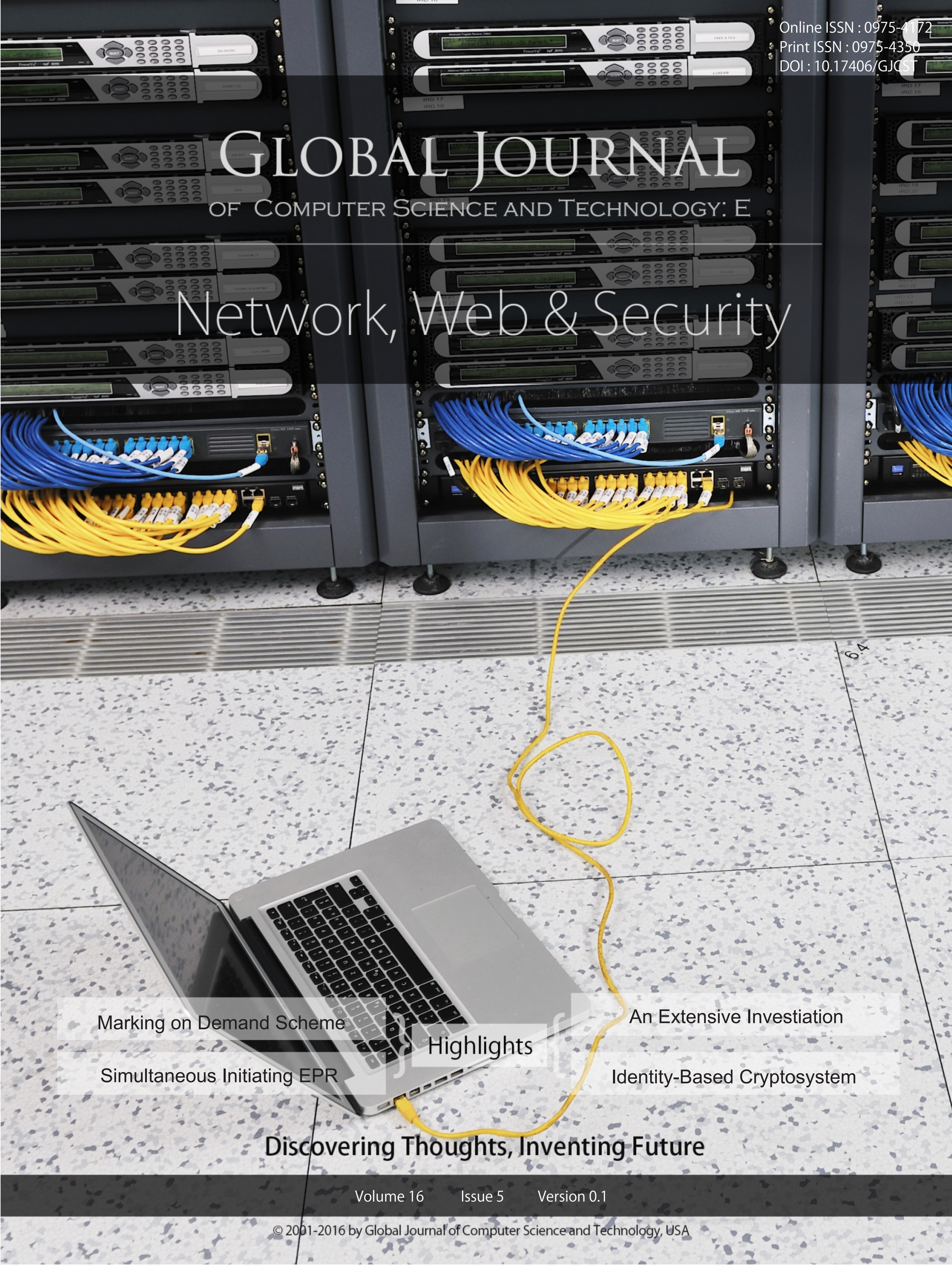
Online ISSN : 0975-4172
Print ISSN : 0975-4350
DOI : 10.17406/GJCST

Global Journal of Computer Science and Technology: E

Network, Web & Security



Marking on Demand Scheme

Simultaneous Initiating EPR

Highlights

An Extensive Investiation

Identity-Based Cryptosystem

Discovering Thoughts, Inventing Future

Volume 16 Issue 5 Version 0.1

Global Journal of Computer Science and Technology: E
Network, Web & Security

Global Journal of Computer Science and Technology: E
Network, Web & Security

Volume 16 Issue 5 (Ver. 1.0)

Open Association of Research Society

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

Global Journals Inc.

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

Sponsors: Open Association of Research Society
Open Scientific Standards

Publisher's Headquarters office

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

Offset Typesetting

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

Packaging & Continental Dispatching

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

Find a correspondence nodal officer near you

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

eContacts

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

Pricing (Including by Air Parcel Charges):

For Authors:

22 USD (B/W) & 50 USD (Color)

Yearly Subscription (Personal & Institutional):

200 USD (B/W) & 250 USD (Color)

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

Contents of the Issue

Global Journal of Computer Science and Technology: E
Network, Web & Security
Volume 16 Issue 5 Version 1.0 Year 2016
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Simultaneous Initiating EPR and Quantum Channel by Quantum Key Distribution Protocol

By Abdulbast Abushgra & Khaled Elleithy
University of Bridgeport

Abstract- Cryptography is the background of protecting the flowed information between various communicated parties. Quantum cryptography gives the extreme trust to transferred information by creating a unique secret key that is based upon the law of physics. This paper will discuss a novel algorithm that is presented through quantum key distribution (QKD) protocol. This QKD protocol depends on parallel quantum communications between participants within EPR and quantum channels. The proposed protocol utilizes the EPR channel to prove the authentication while the quantum channel to transfer the shared key. Moreover, the proposed protocol initiates the verification of the participant's identity between the communicators by the EPR channel. After that the transferred data into quantum channel will create the secret key that contains a string of qubits as well as no need to communicate into classical channel.

Keywords: entangled states, epr pair paradox, intercept-resend attack (IRA), open-key string (OKS), and pauli-matrices measurement.

GJCST-E Classification : C.2.2 D.2.7

SIMULTANEOUSINITIATINGEPRANDQUANTUMCHANNELBYQUANTUMKEYDISTRIBUTIONPROTOCOL

Strictly as per the compliance and regulations of:

Simultaneous Initiating EPR and Quantum Channel by Quantum Key Distribution Protocol

Abdulbast Abushgra [α] & Khaled Elleithy [σ]

Abstract- Cryptography is the background of protecting the flowed information between various communicated parties. Quantum cryptography gives the extreme trust to transferred information by creating a unique secret key that is based upon the law of physics. This paper will discuss a novel algorithm that is presented through quantum key distribution (QKD) protocol. This QKD protocol depends on parallel quantum communications between participants within EPR and quantum channels. The proposed protocol utilizes the EPR channel to prove the authentication while the quantum channel to transfer the shared key. Moreover, the proposed protocol initiates the verification of the participant's identity between the communicators by the EPR channel. After that the transferred data into quantum channel will create the secret key that contains a string of qubits as well as no need to communicate into classical channel.

Keywords: entangled states, epr pair paradox, intercept-resend attack (IRA), open-key string (OKS), and pauli-matrices measurement.

I. Introduction

According to several studies in the quantum cryptography, approving the stability of quantum key distribution protocol (QKDP) is based upon resisting the QKD protocol to quantum security attacks. These attacks have different algorithms and mechanisms that are generally used to tap or eavesdrop transferred data between various parties. The robust scenario in using quantum cryptography is its independency to utilize the law of physics through the quantum channel, which can detect an error as long as it occurs during an eavesdropper or fiber-optics noise. For instance, Intercept-Resend-Attack (IRA) is the well-known quantum attack that threatens the submitted photons from the sender to the receiver (Acín, Masanes, & Gisin, 2003; Curty & Lütkenhaus, 2005). In this scenario, Eve will mask itself as one of the legal parties where she will measure the first particle of the submitted entangled state, and she will try to resend the new created qubit back to Bob. First, the EPR pairs are anticipated to be located with Alice and Bob, but Eve will not be detected at first check. However, because of the property of EPR pairs, Eve will be detected during the second error check that is because EPR pairs have collapsed (Li & Zhang, 2006; Long & Liu, 2002).

The majority of QKD protocols face a difficulty of identity's determination, where the communicators sometimes are not exactly sure who is the sender (or the receiver). Several quantum attacks take this advantage of missed identification between the communicated parties. Therefore, the run time execution will suffer a delay due to much time to restart a new communication or errors correction, every time when the participants find a noise in the quantum channel. On the other hand, the shared data will be lacked if the connected parties ignore the error rate that usually happens during many quantum attacks.

Furthermore, using an authentication procedure at the beginning of the communication between two or more parties will rise the security rate of data transmission. It can also avoid the Intercept-Resend Attack (IRA) or Man-In-Middle Attack (MIM) (Gao, Qin, Guo, & Wen, 2011; Peev et al., 2005) that are based upon impersonating the sender or receiver or both. On the other hand, making a separation between the authentication phase (e.g. EPR channel) and the data submission stage (e.g. Quantum channel) will increase the live time execution that causes a chance for Eve to catch or interrupt even a few communication qubits. Therefore, merging the authentication and the submission of data have the possibility to reduce any eavesdropping chance.

This paper will introduce a new quantum key distribution algorithm, which uses the two quantum channels to ful fil the authentication between the participants by EPR channel. Then the quantum channel will be prepared at the same time of EPR communications to submit a qubits (secret key data). There will be early decision available to both communicators to finish or keep the connection. First part of this paper will demonstrate the initiation of EPR and Quantum channels, and then will show the measurement techniques that will be used at the receiver side.

II. The Initiation of the EPR Connection

In 2015, a quantum key distribution algorithm (Abushgra &Elleithy, 2015) was presented, where it was designed to be robust against common quantum attacks. One such quantum attack was the Man-In-Middle (MIM) attack, which causes an enormous leak of

Author α: He is PhD candidate at Computer Science and Engineering Department, University of Bridgeport.
e-mail: aabushgr@my.bridgeport.edu
Author σ: He is Professor at Computer Science and Engineering Department, University of Bridgeport. e-mail: elleithy@bridgeport.edu

data into the quantum communication between Alice and Bob. The proposed protocol prevents the MIM attack according to the rules of MIM attacks. The MIM attack relies on the fact that the MIM attack will lie or pretend to be a sender or a receiver to both legitimate parties (Yong, Huadeng, Zhaohong, &Jinxiang, 2009). Moreover, the MIM attacker plays on the weaknesses of verification identities between the communicated participants.

The proposed protocol is initiated by a communication into the EPR channel, where Alice (or third party) submits a string of entangled states $|\vdash\Psi\pm\rangle$ or $|\vdash\Phi\pm\rangle$ as well as an unknown state$|\vdash\varphi\rangle$. The unknown state is considered to be the identification state, where the identification state includes initiated strings of time t_1, size of matrix m, and number of matrices n, parity strings p, number of states s, raw index R, and determinate time t_2. The EPR communication will not take a long time of execution because the string of entangled states should be sent in short. After that Bob measures the upcoming string based on EPR theory (Entangled states) (Bell, 1964; Ekert, 1991; Li & Chen, 2007), and then after tensor EPR state (in random) with unknown state (Alice knows) Bob receives a separate code to apply the proper gate, which are one of the quantum gate (x, y, and z gates). Bob will use these gates to measure the states in the superposition. Next, Alice now knows that Bob had received a portion of the right qubits if the percentage of matched qubits is over 70%. Hence, Alice starts

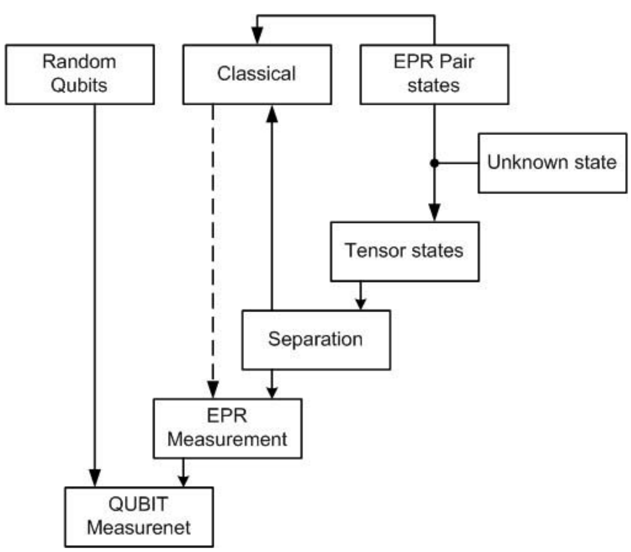


Fig. 2 : Shows the proposed scheme between two legitimate parties (A and B).

The submitted Open-Key (OK) string provides the authentication by EPR entangled states, where each photon is prepared by the sender or third party to be merged with an unknown state (e.g. two dimension state). Measuring an electron at the same time gives an negotiations with Bob to make sure there is no eavesdropper. If Alice finds the matched qubits less than 70%, she will announce Bob to restart another communication.

In case, Alice accepts the EPR communication outcomes, she will submit the string of qubits (data) as in (Abushgra & Elleithy, 2015) into the quantum channel. When Alice initiates the quantum communication within the quantum channel, she knows that Bob has already produced Open-Keys such as (t1, n, m, s, p, R, and t2). On the other side, Bob measures the upcoming qubits based on the number of states (s). He will have enormous amount of measured qubits, where these qubits will be reset in a number of matrices (n) based upon the raw index (R). After that Bob inserts the parity diagonal string (p) into the matrix to start correcting the error phase. If the total of matrix raw summation was even, it means there is no interruption. On the other hand, if the total of the matrix raw was odd, Bob will initiate reconciliation phase.

$$A_{string} = \{t_1,\quad m,\quad n,\quad p,\quad s,\quad R,\quad t_2\}.$$

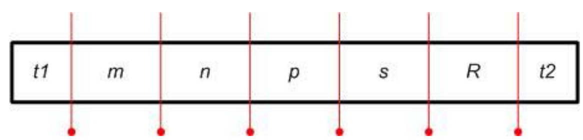


Fig. 1 : Shows the initiated open-key string that will be submitted by Alice to Bob through EPR channel.

opposite result at each participant's side by conservation of linear momentum (Hwang & Lee, 2007). Therefore, these electrons are employed in the authentication phase because physically the photons that represent the Open-Key travel faster than the light speed. Moreover, the Open-Key string in the proposed protocol includes the following characters that are used to authenticate the communication between Alice and Bob as follows:

- t_1 is the initiated time.
- n is the used matrices that can be any number (i = 1, 2,… N).
- m represents the size of the matrix (or matrices) that must be (a = b).
- p is the string of parity diagonal, which it should be prepared simultaneously with EPR connection.
- s is the number of states that are bounded in two types: orthogonal states, or non-orthogonal states.
- R is the row indices sequentially.
- t_2 is termination time.

These characters must be submitted into the Open-Key (OK) string by the EPR channel, and both of the participants should know the included qubits by the theory of entangled states. To measure the upcoming qubits, it is necessary to use the Pauli-Matrices $(\sigma_x, \sigma_y, and \sigma_z)$(Shor & Preskill, 2000) in Bob's circuit's side. Moreover, when Alice desires to share a classical bit 0 with Bob, she initiates the EPR pairs in the state of $|\Phi \mp\rangle$. Also, Alice creates $|\Psi \mp\rangle$ state, if she wants to share classical bit 1(Li & Zhang, 2006).

$$|\varphi\rangle = \alpha|0\rangle + \beta|1\rangle \qquad \textit{Unknown state}$$
$$|\Psi \mp\rangle = \frac{1}{\sqrt{2}}(|0\rangle \otimes |0\rangle \mp |1\rangle \otimes |1\rangle) \quad \textit{Entangled state}$$
$$|\Phi \mp\rangle = \frac{1}{\sqrt{2}}(|0\rangle \otimes |1\rangle \mp |1\rangle \otimes |0\rangle) \quad \textit{Entagled state.}$$

Hence, the submitted particle should be initiated in the previous entangled state, where the position of eigenstate in the $|\emptyset \mp\rangle$ are first $|0\rangle, |1\rangle$ and second $|0\rangle, |1\rangle$. Then Alice keeps one of the qubits in her quantum memory and submits the other qubits into EPR channel. To figure out how the size of the used matrix (or matrices), Bob must calculate the upcoming qubits during the EPR channel in the equation as follows:

$$M_{xy} = \frac{\sum_{i=1} |\varphi_i\rangle}{R} \times n.$$

Based on the received qubits, Bob can organize the qubits into a matrix (or matrices) by the above equation of M_{xy} where the whole received qubits are put in the number of matrices n. Also, the $\sum_{i=1} |\varphi_i\rangle$ is an Open-Key string that represents the tensor of all received qubits. Then Bob begins multiple sequential steps to decide if the qubits are zero eavesdropping or there was a noise during the communication.

III. The Measured Qubits Into EPR Channel

To re-sort the proper indices in their positions, Bob should match the measured indices (R_j) with the OKP (R_i) indices, which usually will be raw by raw. The concluded matrix will be filled in by qubits either $|\emptyset \mp\rangle$ or $|\varphi \mp\rangle$ as well as the diagonal of the matrix (LEFT to RIGHT) that will be filled by a parity string. The parity string (p) is the qubits that should be located at the matrix's diagonal (UP to DOWN). Later, Bob sums the qubits in each raw; if the summation is (0) that means the first correcting phase is secure. Otherwise, Bob will know that there is a noise or an eavesdropping when he finds (1) as a summation of the matrix raw.

$$R_{(i)} = R^*_{(j)}$$

where R is the index number of the matrix, and $i and j \in \{1, 2 \dots n\}$).

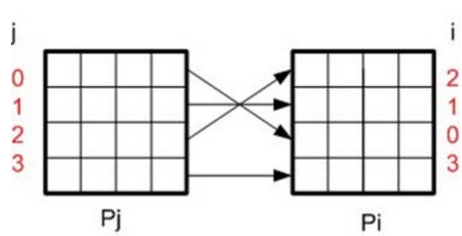


Fig. 3 : Shows re-sorting the received rows by Bob in the proposed protocol between two matrices, where these rows were received such as one string and sequentially resorted in equal matrix.

The abovementioned security checks are not the only security procedures into the proposed protocol, where the implemented decoy states during Alice's preparation is a type of security protection against MIM attacks. The decoy states are located in the upper-triangle of the matrix ($\mu_{ij} \in \{|0\rangle, |1\rangle, |\varphi\rangle, and\ |\emptyset\rangle\}$), where it has a limited tolerance to lose some qubits through the communication phase.

$$\begin{pmatrix} \omega_{11} & \cdots & \mu_{1j} \\ \vdots & \ddots & \vdots \\ \varphi_{i1} & \cdots & \omega_{ii} \end{pmatrix} \times Open - Key = \begin{pmatrix} \delta_{11} & \cdots & \delta_{1j} \\ \vdots & \ddots & \vdots \\ \delta_{i1} & \cdots & \delta_{ii} \end{pmatrix},$$

where $|\varphi_{ij}\rangle$ is the real qubits that will create the key, $|\omega_{ii}\rangle$ is the parity states that are placed diagonally in the matrix, $|\mu_{ij}\rangle$ is decoy states that usually are created similar to real data in random, and $|\delta_{ij}\rangle$ is the resorted matrix's rows after the measurement by Bob ($i \neq j \in \{1, 2 \dots n\}$) as shown in figure (3).

The submitted qubits will not be effected by eavesdroppers, in case, Eve tried to interrupt the channel. The reason of standing against any Eve's interruption is involved through inability of realizing the real qubits of the decoy qubits. Moreover, the string of qubits will be such as one string of data, and there is no variation between each photon.

$\lvert\omega_{ij}\rangle$	$\lvert\varphi_{ij}\rangle$	$\lvert\varphi_{ij}\rangle$	$\lvert\varphi_{ij}\rangle$	$\lvert\varphi_{ij}\rangle$
$\lvert\Psi_{ij}\rangle$	$\lvert\omega_{ij}\rangle$	$\lvert\varphi_{ij}\rangle$	$\lvert\varphi_{ij}\rangle$	$\lvert\varphi_{ij}\rangle$
$\lvert\Phi_{ij}\rangle$	$\lvert\Psi_{ij}\rangle$	$\lvert\omega_{ij}\rangle$	$\lvert\varphi_{ij}\rangle$	$\lvert\varphi_{ij}\rangle$
$\lvert\Psi_{ij}\rangle$	$\lvert\Phi_{ij}\rangle$	$\lvert\Psi_{ij}\rangle$	$\lvert\omega_{ij}\rangle$	$\lvert\varphi_{ij}\rangle$
$\lvert\Phi_{ij}\rangle$	$\lvert\Psi_{ij}\rangle$	$\lvert\Phi_{ij}\rangle$	$\lvert\Phi_{ij}\rangle$	$\lvert\omega_{ij}\rangle$

Fig. 4 : Shows the prepared qubits in one matrix by three classifications, shared data, decoy states, and parity states resorted from up to down and left to right sequentially, where $|\omega\rangle$ is the parity diagonal states, $|\varphi\rangle$ is the data thatwill build the secret key, and $|\mu\rangle$ is the decoy states. ($i = j \in \{1,2 \dots N\}$).

IV. Transferred Qubits Into the Quantum Channel

Alice initiates the qubits that she desires to share with Bob at the same time while preparing the EPR channel. Also, Alice should have the created qubits in her memory to start submitting one by one in a string mode. Although the participants are looking to exchange secure data, the EPR connection, at first, is used to solve the authentication phase. Moreover, both

Global Journal of Computer Science and Technology (E) Volume XVI Issue V Version I Year 2016

parties now attempt to obtain correct data rather than interrupted qubits by the eavesdropper or environment noise. The submitted qubits will be in four states and two non-orthogonal bases.

$$|\varphi\rangle = \frac{1}{\sqrt{2}}(\alpha|0\rangle + \beta|1\rangle),$$

$$|\emptyset\rangle = \frac{1}{\sqrt{2}}(\alpha|0\rangle + \beta|1\rangle).$$

There are multiple options available to transfer a qubit through quantum channel and make the submission secure. One such option is that Alice can communicate with Bob in multi-states $\oplus|s_k\rangle$, where Alice decides through the EPR channel the dimension of the used photon that will be submitted to Bob (e.g. two dimension or more). This is an optional technique that is used; especially, when the secret key should be created to match big data such as in OTP.

Therefore, the proposed algorithm proved its stand against two common quantum attacks. These attacks as mentioned above are IRA and MIM attacks, which both of these attacks are still considered the most concerns around submitting a data through a quantum channel. Also, there is ability to create a huge secret key to match the whole data as long as the quantum memory is available.

V. Conclusion

The proposed QKD algorithm has proved its stability of trusted communication through the quantum channel as well as it is robust against MIM and IRA attacks. The protocol was built, in general, to fulfill the authentication between the communicated parties through the quantum channel. Moreover, the QKD protocol has employed simultaneous exchanges either into the EPR channel (authentication) or quantum channel (sharing a secret key) that maximally sustains the flowing of data into secure phase. As a result, the proposed protocol has been tested and simulated mathematically by MATLAB in classical system and has proved its security against common quantum attacks. Therefore, the proposed protocol is specified by using two parallel quantum channels to prove the authentication between the communicated parties before exchanging secret key plain-text.

References Références Referencias

1. Curty, M., & Lütkenhaus, N. (2005). Intercept-resend attacks in the Bennett-Brassard 1984 quantum-key-distribution protocol with weak coherent pulses. Physical Review A, 71(6), 062301.
2. Ekert, A. K. (1991). Quantum cryptography based on Bell's theorem. Physical Review Letters, 67(6), 661.
3. Gao, F., Qin, S.-J., Guo, F.-Z., & Wen, Q.-Y. (2011). Dense-coding attack on three-party quantum key distribution protocols. Quantum Electronics, IEEE Journal of, 47(5), 630-635.
4. Hwang, T., & Lee, K.-C. (2007). EPR quantum key distribution protocols with potential 100% qubit efficiency. Information Security, IET, 1(1), 43-45.
5. Li, X., & Chen, L. (2007). Quantum authentication protocol using bell state. Paper presented at the Data, Privacy, and E-Commerce, 2007. ISDPE 2007. The First International Symposium on.
6. Li, X., & Zhang, D. (2006). Quantum information authentication using entangled states. Paper presented at the Digital Telecommunications,, 2006. ICDT'06. International Conference on.
7. Long, G.-L., & Liu, X.-S. (2002). Theoretically efficient high-capacity quantum-key-distribution scheme. Physical Review A, 65(3), 032302.
8. Peev, M., Nölle, M., Maurhardt, O., Lorünser, T., Suda, M., Poppe, A., . .Zeilinger, A. (2005). A novel protocol-authentication algorithm ruling out a man-in-the middle attack in quantum cryptography. International Journal of Quantum Information, 3(01), 225-231.
9. Shor, P. W., &Preskill, J. (2000). Simple proof of security of the BB84 quantum key distribution protocol. Physical Review Letters, 85(2), 441.
10. Yong, W., Huadeng, W., Zhaohong, L., &Jinxiang, H. (2009). Man-in-the-Middle Attack on BB84 Protocol and its Defence. Paper presented at the Computer Science and Information Technology, 2009. ICCSIT 2009. 2nd IEEE International Conference on.

Global Journal of Computer Science and Technology: E
Network, Web & Security
Volume 16 Issue 5 Version 1.0 Year 2016
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

An Extensive Investigation on Coronory Heart Disease using Various Neuro Computational Models

By D. Rajeswara Rao & Dr. JVR Murthy
KL University

Abstract- The diagnosis of heart disease at the early time is important to save the life of people as it is absolutely annoying process which requires extent knowledge and rich experience. By and large the expectation of heart infections in conventional method for inspecting reports, for example, Electrocardiogram - ECG, Magnetic Resonance Imaging- MRI, Blood Pressure- BP, Stress tests by medicinal professionals. Presently a-days a huge volume of therapeutic information is accessible in restorative industry in all maladies and these truths goes about as an incredible source in foreseeing the coronary illness by the professionals took after by appropriate ensuing treatment at an early stage can bring about noteworthy life sparing. There are numerous systems in ANN ideas which are likewise contributing themselves in yielding most elevated expectation precision over medical information. As of late, a few programming devices and different techniques have been proposed by analysts for creating powerful decision supportive systems.

More over many new tools and algorithms are continued to develop and representing the old ones day by day. This paper aims the study of such different methods by researchers with high accuracy in predicting the heart diseases and more study should go on to improve the accuracy over predictions of heart diseases using Neuro Computing.

Keywords: *artificial neural networks (ANN), heart diseases, neuro computing.*

GJCST-E Classification : *C.1.3 F.1.1*

An Extensive Investigation on CoronoryHeart Disease using Various Neuro Computational Models

D. Rajeswara Rao [α] & Dr. JVR Murthy [σ]

Abstract - The diagnosis of heart disease at the early time is important to save the life of people as it is absolutely annoying process which requires extent knowledge and rich experience. By and large the expectation of heart infections in conventional method for inspecting reports, for example, Electrocardiogram - ECG, Magnetic Resonance Imaging- MRI, Blood Pressure-BP, Stress tests by medicinal professionals. Presently a-days a huge volume of therapeutic information is accessible in restorative industry in all maladies and these truths goes about as an incredible source in foreseeing the coronary illness by the professionals took after by appropriate ensuing treatment at an early stage can bring about noteworthy life sparing. There are numerous systems in ANN ideas which are likewise contributing themselves in yielding most elevated expectation precision over medical information. As of late, a few programming devices and different techniques have been proposed by analysts for creating powerful decision supportive systems.

More over many new tools and algorithms are continued to develop and representing the old ones day by day. This paper aims the study of such different methods by researchers with high accuracy in predicting the heart diseases and more study should go on to improve the accuracy over predictions of heart diseases using Neuro Computing.

Keywords: artificial neural networks (ANN), heart diseases, neuro computing.

I. Introduction

Heart is most paramount part of the body. Life itself is dependent on the efficient working of the heart. Any realistic predicament in heart has an instantaneous impact on the survival of concerned person that it affects different components of the body. Heart disease is the disease predicated on the performance of the heart. Several factors increases risk of Heart disease like cholesterol, lack of physical exercise, high blood pressure, smoking and exorbitant corpulence. At present, most of the people are suffering from heart disease so there is a need to precise diagnosis at early stages then followed by subsequent treatment that can result the preserving of life. The incipient data relinquished by the National Heart, Lung, and Blood Institute (NHLBI) shows that especially women in the older age groups are in peril of getting heart disease than other people. Recent study withal verbalized that Heart disease can be controlled, if it is diagnosed at an early stage. But it's not facile to control and to do precise diagnosis because of many perplexed factors of heart diseases, like many clinical symptoms are linked with other human organs also, often heart diseases can exhibit sundry syndromes. In-order to scale back the analysis time and to amend the analysis precision, it has end up extra of a inductively sanctioning, to enhance the nontoxic and the puissant clinical determination aid methods to improve the analysis decision approach. Medical diagnosis is intricate process; hence the approach is to develop an accurate system.

Author α: Department of Computer Science and Engineering, KL University, Andhra Pradesh. e-mail: rajeshduvvada@kluniversity.in
Author σ: Professor, Dept of computer science & engineering, JNTU-Kakinada, Andhra Pradesh.

II. Literature Survey

To gain the background knowledge this paper presents a literature survey on neuro computing Techniques for diagnosing Heart disease.

Laercio Brito Gonçalves etal (2016) shown that the Inverted Hierarchical Neuro-Fuzzy Binary house Partitioning which was centred on the Hierarchical Neuro-Fuzzy Binary area Partitioning model (HNFB) that gave an proposal of recursive partitioning to allowed a large number of inputs. The classification method of HNFB-1 has been evaluated with exclusive benchmark databases akin to heart ailment datasets. For interpretable fuzzy ideas it allowed the knowledge extraction [1].

Durairaj M, Revathi V (2015) proposed newly system to obtain more accuracy using back propagation multilayer perceptron (MLP) Algorithm of neural networks than the other neural networks. It is a popular effective method of ANN training network with some optimized techniques like gradient desent where it propagates back to hidden layer. This learning rule moves the network down the steapest slope in error space. The method computes the depth of the loss function in the input data with respect to all the weights in the network. As back propagation algorithm necessitates the activation function as it is applied to multilayer feed forward networks which needs differentiable activation functions. The dataset used for experimentation is Information of heart disease dataset taken from UCI machine learning repository called

Cleveland dataset with 14 attributes, 303 Instances and after cleaning of data they have taken 297 Instances out of 303 Instances. MLP back propagation is trained with the inputs that are adjusted purelin function automatically for increasing the output accuracy. The main aim is to minimize the average sum of errors. The feed forward back propagation algorithm secures highest accuracy of 96.30%. so that the experimental results showed that this algorithm can be effective to predict the heart disease with more accuracy [2].

Noura Ajam (2015) stated that artificial neural networks (ANN) shows the significant results in heart disease diagnosis. The architecture of neural network is formed by number of processing units (Neurons) and connections between them. A subgroup of processing elements is called layer. The number of neurons and the layers depends upon the complexity of the system studied. ANN is widely used in medical diagnosis and health care applications because of their predictive powerful classifier for tasks, fault tolerance, generalization and learning from environment. ANN is unsupervised learning type provided only with inputs, but no known targets. It is self organized. The Dataset used is Cleveland dataset which consists of 14 attributes and 303 instances. ANN is trained using back propagation learning algorithm on heart disease data. Input and target samples are divided as 60% training set, 20% validation set, 20% test set. The activation function of tangent sigmoid for hidden layers and linear transfer function for output layer is used. Mean square error "MSE" is calculated which is equal to 0.1071 and the classification accuracy for heart disease is 88% [3].

S. Florence, N.G. Bhuvaneswari Amma, G.Annapoorani, K.Malathi (2014) proposed the system which uses the Neural Network and the Decision Tree (ID3) for the prediction of heart Attacks. The dataset used is acath Heart attack dataset provided by the UCI machine Learning Repository which has attributes are considered to diagnose the heart attacks. CART , ID3, C4.5 decision tree algorithms used Gini index to measure the impurity of a partition or set of training attributes. The dataset contains 6 attributes like age, sex, cardiac duration, signal, possibility of attack. The final one is the class label. Depends upon the attribute values present in the dataset the corresponding class label that is the prediction is happening at the final stage. For training and testing the network where 75 percent is used for training and 25% is used for testing the system. The knowledge obtained from the classification is used to test the system. When comes to the neural network the input layer has 6 nodes, the hidden layer has 3 nodes and the output layer consists of 2 nodes. Finally it shows 2 outputs, that is the possibility of heart attacks. The prediction is done using the tool called Rapid Miner Studio. Results have been generated by using the Decision Tree as well as the Neural networks. The graphs have been plotted and these are generated in a simulator called Rapid Minor Tool. They had just drawn the results and have predicted whether is there is an attack or not using these Networks [4].

Hlaudi Daniel Masethe etal (2014) discussed data mining classification algorithms for predicting the heart attacks are J48, Naïve Bayes, REPTREE, CART and Bayes. The aim is to predict possible heart attacks using data mining techniques from the patient dataset and determines the best model which gives the correct predictions of highest percentage for diagnoses. From medical practitioners, the dataset of eleven attributes are collected for the predictions of heart disease. The attributes are labelled as Patient Id-Number, Gender of the patient, Cardiogram report, Age of the patient, Chest Pain type, Blood Pressure Levels, Heart Rate, Cholesterol levels, Smoking habitat, Alcohol consumption and levels of Blood Sugar. For the prediction, the weka data mining tool is used to analysis to discover the patterns. The algorithms have been applied on the data set and thus the results have been obtained and they had observed that the J48, REPTREE and SIMPLE CART show a prediction model of 89 cases with a risk factor positive for heart attacks. The best classification technique to be J48, REPTREE and SIMPLE CART algorithm perform similar in this data set, while Bayes Net algorithm out-performed the other techniques. Thus the algorithms results do not show any difference in the prediction when using different classification algorithms in data mining [5].

Jayshril S. Sonawane, D. R. Patil (2014) evaluated the prediction approach for coronary heart sickness making use of studying LVQNN algorithm. The neural network on this algorithm had 13 attributes of input and predicts the presence or absence of heart ailment of sufferer. The prediction method is based on ANN. Synthetic neural network is an understanding processing procedure that strategies the expertise in an identical method because the biological apprehensive system techniques. On this technique neural network considers that they've got competencies to derive that means from tricky or imprecise capabilities which possibly used to extract specified patterns and discover trends which can be extra elaborate to be seen through both humans or other pc techniques and approaches. The essential cause of utilizing LVQ is that it creates prototypes which can be effortless to interpret for specialists within the respective utility area. Learning vector Quantization is aggressive network uses the supervised learning methodology which contains two layers specifically competitive layer and linear layer. The learning vector quantization algorithm is applied on the Cleveland heart ailment database. This suggests that the prediction approach offers higher performance consequently giving us an effective method for the prediction for the heart sickness [6].

Jesmin Nahar and Tasadduq Imam etal (2013) examined the actual fact of computational smart techniques in coronary disorder discovering. Cleveland knowledge was once used to participate in connection with six comprehended classifiers. For most classifiers and higher section knowledge set the execution was once elevated by way of encouraged feature decision. They developed an efficient algorithm for rule extraction test on coronary heart ailment information for various associative rule mining algorithms such as Apriori, Predictive Apriori and Tertius for the analysis of coronary heart diseases diagnosis [7].

Sanjeev Kumar, Gursimranjeet Kaur (2013) detected the heart ailments in individual by means of utilising the Fuzzy expert system. The designed approach is based on two hospitals dataset and international lab database. Comparative analysis is finished between these two hospitals dataset and the lab database methods. Via utilising the fuzzy knowledgeable approach the diagnosis of heart disease has been carried out which consists of six inputs and a pair of outputs. The six attributes are chest anguish, Blood strain, ldl cholesterol, Blood Sugar, Max coronary heart expense and old peak. Rule base is the important phase in fuzzy inference method and first-rate of results in a fuzzy process depends upon the fuzzy principles. These principles have been applied on enter variables to verify how effective the fuzzy approach works. Using the informed method the established results for prognosis of coronary heart disorder have the foundations that if the worth is low then the danger is low if the value is excessive then the hazard is excessive. For this reason the trained procedure has been carried out and suggests that it is extra effective for analysis of heart ailment [8].

Dhanashree S. Medhekar, Mayur P. Bote, Shruti D. Deshmukh (2013) presented a classifier technique for the heart sickness prediction and likewise they've confirmed how the Naïve Bayes can be used for the classification cause. They will categorise clinical knowledge to five distinct classes namely no, low, normal, excessive, very excessive. If discovered any unknown sample the method will classify into its respective class label of that sample. The dataset used here is the Cleveland medical institution groundwork coronary heart disease set which contains 303 observations and 14 parameters. The system works in two phases: coaching phase, testing phase. In the training segment the classification is supervised, classifies knowledge situated on the training set and sophistication labels as a classifying attribute and classifies into new knowledge. In the checking out segment it involves the prediction of the unknown knowledge or the lacking values. The Naïve Bayes algorithm is used and it is situated upon the Bayesian theorem. The outcome has proven that the accuracy has been obtained through altering the number of occasions within the given dataset [9] .

Akin Ozcift and Arif Gulten (2011) developed a Random forest "RF" ensemble classifier to assess their classification of performances utilizing Parkinson's, diabetes and coronary heart diseases data sets. Using correlation situated characteristic determination algorithm three knowledge sets were minimized after which performances of 30 machine learning algorithms were estimated for three information sets and constructed situated on RF algorithm [10].

Mai Shouman, Tim Turner, Rob Stocker (2011) have applied a wide range of techniques to different types of Decision tree seeking better performance in diagnosing the heart disease. They have proposed a model that outperforms J4.8 Decision tree and bagging algorithm in diagnosing heart disease patients. They have proposed a model that involves different discretization techniques, multiple classifiers voting technique and different Decision tree type for diagnosis of heart patients. Different combinations of discretization methods, decision tree types and voting are tested to identify which combination will provide the best performance in diagnosing heart patients. Data discretization is divided into supervised and unsupervised methods. The unsupervised methods involve equal width and equal frequency while the supervised discretization methods involve chi merge and entropy. The data partitioning involves testing with and without voting. Three Decision Tree types are tested: Information Gain, Gini Index, and Gain Ratio. Finally, reduced error pruning is applied on all the Decision Tree rules extracted from the training data. The Dataset used is the Cleveland Clinic Foundation heart disease consists of 76 raw attributes. The results show us that highest accuracy is obtained by the equal width discretization Information Gain Decision Tree with 79.1%. Different partitions of voting were applied to the data. The highest accuracy achieved by the equal frequency discretization in Gain Ratio Decision Tree is 84.1%. When compared with the existing system this model has shown the best results and has achieved highest accuracy [11].

Shashikant Ghumbre, ChetanPatil, Ashok Ghatol (2011) developed a decision help approach using RBF and SVM. RBF networks are beneficial for continuous or piecewise continuous actual-valued mapping approximations. Three parameters particularly the quantity of basis capabilities, their place and their width determine the measure of accuracy of the RBF networks. SVM is a class of common feed ahead networks like Radial-groundwork perform networks. SVM can be utilized for pattern classification and nonlinear regression. Extra precisely, support vector computer is an approximate implementation of the system of structural risk minimization. This principle is based on the actual fact the error expense of a learning computer

Global Journal of Computer Science and Technology (E) Volume XVI Issue V Version I Year 2016
7

on scan information is bounded with the aid of the sum of the educational-error price and term that will depend on the Vapnik- Chervonenkis (VC) dimension. During experimentation, it's observed that, proper and complete data assortment process is the right route for the choice of high-quality classifier. For evaluating generalization performance with appreciate to accuracy, sensitivity, and specificity dataset is partitioned into quantity of subsets (i.E train set and test set). Overall natural performance will depend on accuracy of SVM and RBF utilizing subsets of coaching and test sets. SVM offers highest accuracy with increase in measurement of training data, while RBF gives minimal accuracy with scale down in dimension of test information [12].

Pasi Luukka and Jouni Lampinen (2010)have carried out classification manner situated on pre-processing the info with principal aspect evaluation (PCA) and then utilising differential evolution classifier to the prognosis of coronary coronary heart ailment. This system used to be utilized for predicting prognosis from clinical data units. The outcomes indicated that pre-processing the info before classification would not simplest help with the curse of increasing information dimensionality, but additionally furnish one more enhancement in classification accuracy [13].

Nazri Mohd Nawi etal (2010) have proposed a novel technique to increase the effectiveness of back propagation neural network. In Gradient Descent with Momentum and Adaptive gain proposed calculation, for every hub the addition high-quality used to be modified adaptively to alter initial search. The coronary health problem of the sufferer was predicted productively and the calculations had been firmly developed and can upgrade the computational productiveness [14].

Resul Das and Ibrahim Turkoglu et al (2009) have encouraged unique tools and quite a lot of methodologies to create powerful scientific decision supportive network. A framework used to be offered which makes utilization of Statically analysis procedure (SAS) base programming for diagnosing of the coronary ailment [15].

[16] Hongmei Yan and Jun Zheng (2008) have provided a exact coded GA established framework to decide on the elemental medicinal accessories key to the coronary heart sicknesses choice. It has been proposed to prefer the basic elements and aid the finding of 5 principle heart infections which have been hypertension, coronary health problem, rheumatic valvular coronary sickness, perpetual pulmonale and innate coronary illness.

Kemal Polat and Salih Gunes (2007) offered a hybrid method on medical diagnosis using feature decision, fuzzy weighted pre-processing and artificial Immune Recognition System. The hybrid method have two stages. The datasets of heart ailment and hepatitis disease have been reduced to 9 from 13 & 19 within the feature selection by means of C4. 5 decision tree algorithm. The heart sickness and hepatitis sickness datasets are utilized from UCI database as clinical dataset [17].

Other types of methods which are widely employed in the diagnosis of Heart disease are: Hongmei Yan etal [18] developed a multilayer perceptron based decision support system to support the diagnosis of heart disease. A.T.Sayad etal [19] employed the Multi-layer Perceptron Neural Network with Back-propagation as the training algorithm on heart disease diagnostic system. Resul Das etal [20] introduced a methodology which uses SAS base software for diagnosing of the heart disease. Sunila Godara etal [21] presented a decision support system based on MLP neural network architecture for diagnosing heart disease.

Summary

Table 1 : Summarized table shows proposed models for diagnosing the heart disease

S.NO	Author	PROPOSED MODEL	EXPECTED RESULTS *ACCURACY*
1.	Hlaudi Daniel Masethe etal [5]	Data Mining Techniques	Depends upon conditions provided.
2.	Mai Shouman etal [11]	Nine subsets voting model	79.1%-Info gain 84.1%-Gain ratio
3.	Durairaj M	MLP + Back propagation	Highest accuracy of 96.30%
4.	S.Florence etal [4]	Neural Network+ Decision Tree	Depends upon conditions provided
5.	Sanjeev Kumaretal [8]	Fuzzy Expert System	Risk Factors 0-low,1-high
6.	Dhanashree S. Medhekar etal [9]	Naïve Bayes	Based upon different instances- 89.98%
7.	Shashikant Ghumbreetal [12]	RBF + SVM	Depends upon size of training data
8.	Jayshril S. Sonawanel	LVQ+ Neural Network	85.55%-highest accuracy

	etal[6]		
9.	Sumit Bhatiaetal[22]	SVM + GA	accuracy of 90.57%
10.	Vidyullatha.p, D.Rajeswara Rao[23]	Rough Set Model	good clarity and more Accuracy over the incomplete data set.

III. Conclusion

This learn applied a literature survey of comparative studies on neural networks, machine learning procedures and statistical techniques used for prediction and classification intent of the heart sickness. These evaluation facets out the knowledge of neural networks being employed for classification and prediction of heart attack. In this regard, these systems end up complementary approaches for model constructing rather of competing approaches.

References Références Referencias

1. Laercio Brito Gonçalves and Marley Maria Bernardes Rebuzzi Vellasco.Inverted Hierarchical Neuro-Fuzzy BSP System: A Novel Neuro-Fuzzy Model for Pattern Classification and Rule Extraction in Databases.Journal of IEEE Transactions On Systems, Man, And Cybernetics2016 March; Vol. 36(2)
2. Durairaj M, Revathi V, “Prediction Of Heart Disease Using Back Propagation Mlp Algorithm”, International Journal Of Scientific & Technology Research, Issn 2277-8616 ,Volume 4, Issue 08, August 2015
3. Noura Ajam, “Heart Diseases Diagnoses Using Artificial Neural Network”, Network And Complex Systems, ISSN 2224-610X (Paper) ISSN 2225-0603 (Online).Vol.5, No.4, 2015
4. S.Florence, N.G.Bhuvaneswari Amma, G.Annapoorani, K.Malathi, “Predicting The Risk Of Heart Attacks Using Neural Network And Decision Tree”, International Journal Of Innovative Research In Computer And Communication Engineering, ISSN (Online) : 2320-9801, Vol. 2, Issue 11, November 2014.
5. Hlaudi Daniel Masethe, Mosima Anna Masethe, “Prediction of Heart Disease using Classification Algorithms”, Proceedings of the World Congress on Engineering and Computer Science 2014 Vol II WCECS 2014, 22- 24 October, 2014, San Francisco, USA.
6. Jayshril S. Sonawanel, D. R. Patil, “Prediction of Heart Disease Using Learning Vector Quantization Algorithm”, 978-1-4799-3064-7/14/$31.2014 IEEE.
7. Jesmin Nahar and Tasadduq Imam. Association rule mining to detect factors which contribute to heart disease in males and females.Journal of Expert Systems with Applications. 2013Vol.40:1086–1093.
8. Sanjeev Kumar, Gursimranjeet Kaur, “Detection Of Heart Diseases Using Fuzzy Logic”, International Journal Of Engineering Trends And Technology (IJETT), ISSN: 2231-5381,Volume 4 Issue 6- June 2013, Pp:2694-2699
9. Dhanashree S. Medhekar, Mayur P. Bote, Shruti D. Deshmukh, “Heart Disease Prediction System Using Naive Bayes”, International Journal Of Enhanced Research In Science Technology & Engineering, Issn No: 2319-7463, Vol. 2 Issue 3, March.-2013
10. Akin Ozcift and Arif Gulten, “Classifier ensemble construction with rotation forest to improve medical diagnosis performance of machine learning algorithms.Journal of Computer Methods and Programs in Biomedicine2011:Vol.104:443-451.
11. Mai Shouman, Tim Turner, Rob Stocker, “Using Decision Tree for Diagnosing Heart Disease Patients”, Proceedings of the 9-th Australasian Data Mining Conference (AusDM'11), Ballarat, Australia, Conferences in Research and Practice in Information Technology (CRPIT), Vol. 121.
12. Shashikant Ghumbre, Chetan Patil, And Ashok Ghatol, “Heart Disease Diagnosis Using Support Vector Machine”, International Conference On Computer Science And Information Technology (ICCSIT'2011) Pattaya Dec. 2011
13. Pasi Luukka and Jouni Lampinen A Classification Method Based on Principal Component Analysis and Differential Evolution Algorithm Applied for Prediction Diagnosis from Clinical EMR Heart Data Sets.Journal of Computer Intelligence in Optimization Adaption, Learning and Optimization. 2010;Volume 7: 263-283.
14. Nazri Mohd Nawi and Rozaida Ghazali.The Development of Improved Back-Propagation Neural Networks Algorithm for Predicting Patients with Heart DiseaseIn proceedings of the first international conference ICICA.2010;Vol.6377:317-324.
15. Resul Das and Ibrahim Turkoglu, “Effective diagnosis of heart disease through neural networks ensembles”, Journal of expert system with applications.2009:Vol.36, :7675–768.0
16. Hongmei Yan and Jun Zheng, “Selecting critical clinical features for heart diseases diagnosis with a real-coded genetic algorithmJournal of Applied Soft Computing 2008 :Vol.8:1105-1111.

17. Kemal Polat and Salih Gu nes.A hybrid approach to medical decision support systems: Combining feature selection, fuzzy weighted pre-processing and AIRS. Journal of Computer Methods and Programs in Biomedicine.2007;Vol.88:164-174
18. Hongmei Yan, Yingtao Jiang, Jun Zheng, Chenglin Peng, Qinghui Li," A multilayer perceptron-based medical decision support system for heart disease diagnosis", Expert Systems with Applications 30 (2006) 272–281.
19. T. Sayad, P. P. Halkarnikar," Diagnosis of Heart Disease Using Neural Network Approach", International Journal of Advances in Science Engineering and Technology, ISSN: 2321- 9009, Volume- 2, Issue-3, July-2014,pp:88-92.
20. Resul Das, Ibrahim Turkoglu , Abdulkadir Sengur," Effective diagnosis of heart disease through neural networks ensembles", Expert Systems with Applications 36 (2009) ,ISSN:7675– 7680,pp:7676-7680.
21. Sunila Godara, Nirmal," Yanwei, X et.al [7] developed data mining algorithms for predicting survival of CHD patients. Intelligent and Effective Decision Support System Using Multilayer Perceptron", International Journal of Engineering Research and Applications (IJERA), ISSN: 2248-9622, Vol. 1, Issue 3, pp.513-518.
22. Sumit Bhatia, Praveen Prakash, and G.N. Pillai, "SVM Based Decision Support System for Heart Disease Classification with Integer-Coded Genetic Algorithm to Select Critical Features", Proceedings of the World Congress on Engineering and Computer Science 2008 WCECS 2008, October 22 - 24, 2008, San Francisco, USA.
23. Vidyullatha Pellakuri, D.RRao, "Progressive Decision Making In The Department Of Cardiology By Optimized Rough Set Model", Int J Pharm Bio Sci 2016 April; 7(2): (B) 658 – 665.

Global Journal of Computer Science and Technology: E
Network, Web & Security
Volume 16 Issue 5 Version 1.0 Year 2016
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Review of Contemporary Literature on Machine Learning based Malware Analysis and Detection Strategies

By G.Bala Krishna, Dr. V.Radha & Dr. K. Venu Gopala Rao
KMIT/JNTUH

Abstract- Malicious software also known as malware are the critical security threat experienced by the current ear of internet and computer system users. The malwares can morph to access or control the system level operations in multiple dimensions. The traditional malware detection strategies detects by signatures, which are not capable to notify the unknown malwares. The machine learning models learns from the behavioral patterns of the existing malwares and attempts to notify the malwares with similar behavioral patterns, hence these strategies often succeeds to notify even about unknown malwares. This manuscript explored the detailed review of machine learning based malware detection strategies found in contemporary literature.

Keywords: malware detection, malware signature, API call sequence, anomalies, static analysis, dynamic analysis, machine learning.

GJCST-E Classification : C.2.0 D.4.6 H.2.7

REVIEW OF CONTEMPORARYLITERATUREONMACHINELEARNINGBASEDMALWAREANALYSISANDDETECTIONSTRATEGIES

Strictly as per the compliance and regulations of:

Review of Contemporary Literature on Machine Learning based Malware Analysis and Detection Strategies

G.Bala Krishna [α], Dr. V.Radha [σ] & Dr. K. Venu Gopala Rao [ρ]

Abstract- Malicious software also known as malware are the critical security threat experienced by the current ear of internet and computer system users. The malwares can morph to access or control the system level operations in multiple dimensions. The traditional malware detection strategies detects by signatures, which are not capable to notify the unknown malwares. The machine learning models learns from the behavioral patterns of the existing malwares and attempts to notify the malwares with similar behavioral patterns, hence these strategies often succeeds to notify even about unknown malwares. This manuscript explored the detailed review of machine learning based malware detection strategies found in contemporary literature.

Keywords: malware detection, malware signature, API call sequence, anomalies, static analysis, dynamic analysis, machine learning.

I. Introduction

The term "Malware" stands for malicious software, and it usually specifies as hostile software application. According to G. Mc Graw et al., [1] there are multiple causes as code added, changed, or removed from the software it get corrupt and it deliberately causes harm and disrupt normal computing activity. A virus had a broad range of destructive software applications such as viruses, Trojans, Spywares and other intrusive code [2].

The malware can discriminate by the capability of replication, propagation, self-execution and corruption of the operating system. If the computer system gets extortion it influence on confidential information, integrity and denial of assistance. In malware Replication is a crucial component as it assure its existence.

In some cases Replication generates consumption and continuation of system resources (e. g. hard disk, RAM). If confidential assets are being used by any other malware types other than the user, to conceal themselves from anti-malware detector they use a technique called polymorphic or metamorphic techniques.

Author α: Assistant professor, Dept. of IT, KMIT Hyderabad. e-mail: govind.krishna83@gmail.com
Author σ: Assistant Professor, IDRBT Hyd e-mail: vradha@idrbt.ac.in
Author ρ: Professor, Dept. of CSE, GNITS Hyderabad. e-mail: kvgrao1234@gmail.com

The operating system gets corrupt through data transfer from desecrate device to another protected device familiar, such as executable files, boot records of disk drives or exhausting network bandwidth, by using local or network files system. In such case malware makes operating system susceptibility and few software bugs are faults and it starts its life cycle at the same system and infected system simultaneously by remotely control.

According to a McAfee simplified report (year 2013) says that "malware continues to grow" [3] and by G Data and king [4] [5] soft Laboratory declare n-number of innovative malware will emerge promptly and to build an anti-malware the analyzers and constructors are enhanced by their unique techniques and methods [6]-[10]. To construct a malicious software the techniques which are been used to categorized and estimate in groups such as obfuscation techniques, invocation methods, platform, spreading and propagation techniques.

To actuate a program has a malicious attentive or not, malware detection system is used. In this detection system there are two different functions, detection and analysis [12]. Detection system is a protecting one as it may or may not be prevail in the same system [13] and the tasks can be split into client and server as it analogous in cloud-based antivirus [8, 12]. A numerous renovations are made on detection and analysis functions [5], [12], [15]-[19].

In malware detection system specialized solutions are added to expansion in success and achievement. Such as cloud computing [10], network based detection system [20], web, virtual machine [21], [22], agent technology [23]-[29] or by the use of hybrid methods and technologies.

II. Review of Contemporary Literature and Contributions

In earlier stage malware had come up with signature based detection. But now in this stage malware signature has introduced an automatic generation and it is pretended to be important and it increased its pattern in similar speed.

The signature based detection system has some imperfection as follows to continue the updates of signature it requires high maintenance cost. By inclusion such methods it could be evenly avoided by malware in polymorphic form [30]. To conquer the imperfection, it embraces code in normal vision to grab the consecrate original maliciousness. To vary the polymorphic techniques and apply, this grabbed malicious code is used but still it is anemic to detecting obfuscated malware. Apart from this some execution paths can be explored execution [31] [30].

Due to certain requirements the malware analysis is all ways conserved the techniques in the prior, consequently dynamic analysis was considered. To identify and execute a complicate malware dynamic analysis methods are used. In dynamic analysis the malware shows how it operates and recognize the unknown malware which is identically operates like a known malware [32]. There are two familiar primary dynamic methods are control flow analysis and API call analysis. [33] [34].

API call data display how the malware gets operates and it can be obtained by both static and dynamic approaches. The API list and PE format of the executable files can be derived by the static approach [35] [36] [37] [38]. In dynamic approach. [39] [40] [41], [42] [43] [44] API calls can be recognized by running executable files it usually run in virtual machine.

In API call there two familiar ways to evaluate the data accumulated by static approach. The first one implements simple statistical analysis, for example, to count the frequency of API call which is aspect to organize malware [35]. The second approach is to gather the API call data through data mining or machine learning techniques. In another way the API call sequence data which gathered by the dynamic approach are helpful to creates a behavioral patterns. The information accumulated by the dynamic approach also operates simple statistics such as frequency counting [39] and data mining or machine learning [40] [42] [44].

In other way, researchers are analyzed more ways to develop API call sequence information. In earlier research API call had introduced API call graph [45] with various kinds of call graph analysis. To get more consequential features for call graph analysis, the analyzer had espoused the mechanism of social network analysis. [46] According to analyzers the affinity among API call sequences is based on cosine similarity function and lengthy jaccard measure. Due to modern research [33] [34] [47] [48] more information had been added such as control flow information and API argument information to inflate the efficiency in the mining process.

The API call analysis been done with API call approaches. In this abstraction the dynamic method is applied to excerpt API call sequences. To obtain austerity patterns, DNA sequence alignment algorithms (MSA and LCS) are adapted. With API call sequence patterns and the critical API call sequence, we can recognize the unknown malware or its variation with elevated efficiency.

Anderson et al. [49] defined a malware detection strategy that builds a set of graphs from the given instruction set and then analyzes these graphs to notify the proneness of the malware activity. In order to build the graphs the markov chains were defined on 2-gram sequences. The graphs defined form the training set further used to build a similarity matrix using graph kernel. The graph kernel is the mix of Gaussian and spectral kernels, which are in use to assess the similarity between graph edges and similarity between graphs respectively. Further the support vector machine that learns from the similarities between graph edges and graphs is used to classify the input call sequences.

By using such liberal malware software the multiple kernel is achieved and learning design used in this work to exhibit selective refined differences occurrence of malware. The inadequacy of this approach is computed consequence is very high, hence the use of this approach is discouraged.

Bayer et al. [50] prospect a technique that groupsthe call sequences generated by Anubis [51]. The behavior adequately of the call sequences is considered as objective to cluster the call sequences by Locality Sensitivity Hashing (LSH) [52]. The constraint of the model is that LSH is capable to generate probabilistic clusters.

Biley et al. [53] argued that malware prototyping is not consistent among the notable antivirus products available. In order to this the authors devised a novel classification strategy that classifies the malware according to the changes observed at system state. A strategy that prototypes the behavior of the malware is used, further the malwares are classified according to these behavior prototypes. The distance between a class and a malware is assessed by the distance metric called “normalized compression distance (NCD)”. The constraint observed in empirical study of this model is that the behavior prototype definition is static and limited to malwares that are not fall in zero-day category (unknown malwares). park et al. [54] defined a classification strategy that classifies malware based on the graphs generated from the call sequences. Further graph similarities between confirmed malware call sequence graph and unknown call sequence graph will be assessed. The similarity index is the “max number of subgraphs identified in both graphs”. The malwares those controls the system privileges without initiating the system call sequences are not traceable by this classification model, which is a significant constraint of this model. Firdausi et al. [55] defined a machine learning model for malware detection. The said model initiates the process by exploring the behavioral patterns of the malware samples given for training, which is done by the model called Anubis [51]. Further these observed behavioral patterns will be organized as sparse vectors and learns the behavior prototypes. The malware samples given for testing will be classified, which is based on the behavioral prototypes learned in training phase. The performance of the model is estimated through benchmark classifiers and they are "j48", "multilayer perception neural networks (MLP)" "Naïve Bayes", "Support Vector Machine (SVM)" and "k- Nearest Neighbors (kNN)". The experimental results indicating that the J48 classification delivered much classification accuracy.

Nari et al. [56] devised a network flow behavioral analysis framework for malware detection. The network transactions obtained from PCAP files were considered to extract the network flows. Further a network activity representation graph is drawn from these network flows. The given network flows labeled as malware were used in training phase. Further this framework learns representation of the features such as size of the graph, average, maximum and root level out-degree and count of specific nodes of the network activity graphs of the given input network flows. Further these features specific information uses to classify the input malware samples in testing phase. In order to perform the classification, the WEKA library [57] was used. The experimental study indicating that the J48 is the best classifier among all classifiers available in WEKA library.

Lee et al. [58] explored a machine learning based malwares clustering. The training phase builds the behavioral profiles of the malware samples given as training data and the profile includes the system resources invoked by the system calls and their arguments. Further the similarities between behavioral profiles were considered distance function to cluster the malwares, which was done by k-medoids. The outliers are adjusted to the clusters based on the nearest neighbor strategy. This approach is the combination of static and dynamic clustering strategy that clusters known features by k-medoid and unknown and new features by nearest neighbor approach. This strategy is evincing that hybrid approach is more robust in order to classify the known as well as unknown features effectively.

Another hybrid approach for malware detection was devised by Santos et al. [59]. This approach tracks the known features (static features) through the analysis of the sequence of operational codes in given malicious executable and the unknown features (dynamic features) were noticed from the observation of exceptions and operations in system calls. The experimental study was done under various classifiers and results obtained were evincing the significant accuracy in malware classification Islam et al. [60] explored a similar strategy that extracts static and dynamic features to classify the executables into malevolent or benevolent. The features such as function length, function executable frequency and length of the strings involved are included in known features and the features such as function identity and function arguments are included in unknown features. The experiments were done using the classifiers called Support Vector Machine, Decision Tree and Random Forest and results evincing that the random forest is the best classifier among all considered.

The malware classification method devised by Anderson et al. [61] is using the divergent input sources such as control flow graphs, static call sequences, portioned executables, dynamic call sequences and file signatures. Further this model learns the weight of these input combinations from the given training set of malevolent and benevolent executables. The observed weights of these input combinations are used further to classify the malevolent and benevolent executables during testing phase. The process overhead is the significant constraint of this model observed against dense and high speed network streams.

III. Conclusion

The current era of internet and computer systems are prone to serious security threats due to the malicious software which are also referred as malware. Hence the significant research contributions aimed to define malware detection and prevention strategies in contemporary literature. All of these contributions are fall in the categories of either anomaly based, signature based or call sequence analysis based detection. The signature based models are capable to notify and prevent the malwares that are notified earlier. In contrast to this the anomaly models and call sequence analysis models are capable to identify the malwares based on the similarities learned from previous malware attacks. The difference between the anomaly and call sequence analysis models is that the anomaly based learning models can adopt user defined features, whereas the call sequence analysis models notify the similarities learned from the call sequences of 2-gram, 3-gram or n-gram. This manuscript aimed to affirm the objectives and limits of the contributions found in recent literature. The conclusion of the review evincing that the machine learning based models that learns from either anomalies or call sequence are tolerable the constraints observed in signature based malware detection strategies. The anomaly and call sequence learning models found in contemporary literature are not adequate to defend the challenges evincing from the vibrant and unjust network data. Hence the significant contributions are in demand to handle the challenges evinced in current era of internet and computer system usage.

References Références Referencias

1. McGraw, G. and G. Morrisett, Attacking Malicious Code: A Report to the Infosec Research Council. IEEE Softw., 2000. 17 (5): p. 33-41.
2. Xufang, L., P. K. K. Loh, and F. Tan. Mechanisms of Polymorphic and Metamorphic Viruses. in Intelligence and Security Informatics Conference (EISIC), 2011 European. 2011.
3. Mcafee and Lab, 2013 Threats Predictions. 2013.
4. Berkenkopf, R. B. S., G-Data Malware Report. 2010.
5. Ye, Y., et al., Intelligent file scoring system for malware detection from the gray list, in Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining. 2009, ACM: Paris, France. p. 1385-1394.
6. Rieck, K., Malheur A novel tool for malware analysis 2012.
7. Pinz, C. I., et al., Improving the security level of the FUSION@ multi-agent architecture. Expert Syst. Appl., 2012. 39 (8): p. 7536- 7545.
8. Ammar Ahmed E. Elhadi, M. A. Maarof, and A. H. Osman, Malware Detection Based on Hybrid Signature Behaviour Application Programming Interface Call Graph. American Journal of Applied Sciences, 2012. 9 (3): p. 283-288.
9. Kevadia Kaushal, P. S., Nilesh Prajapati, Metamorphic Malware Detection Using Statistical Analysis. International Journal of Soft Computing and Engineering (IJSCE), 2012. 2 (3).
10. Yanfang Ye, T. L., Shenghuo Zhu, Weiwei Zhuang, EgemenTas, Umesh Gupta, MelihAbdulhayoglu, Combining file content and file relations for cloud based malware detection, in Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. 2011, ACM: San Diego, California, USA. p. 222- 230.
11. Christodorescu, M., et al., Semantics-Aware Malware Detection, in Proceedings of the 2005 IEEE Symposium on Security and Privacy. 2005, IEEE Computer Society. p. 32- 46.
12. Yin, H., et al., Panorama: capturing systemwide information flow for malware detection and analysis, in Proceedings of the 14th ACM conference on Computer and communications security. 2007, ACM: Alexandria, Virginia, USA. p. 116-127.
13. Vinod, P., et al., Survey on Malware Detection Methods. 2009.
14. Zeltser, L., what is cloud Anti-Virus and how it does work.
15. Jiang, X., X. Wang, and D. Xu, Stealthy malware detection through vmm-based "outof- the-box" semantic view reconstruction, in Computer and communications security. 2007, ACM: Alexandria, Virginia, USA. p. 128-138.
16. Automated dynamic binary analysis. 2007.
17. Deepak Venugopal, G. H., Efficient signature based malware detection on mobile devices. Mob. Inf. Syst., 2008. 4 (1): p. 33-49.
18. Kolbitsch, C., et al., Effective and efficient malware detection at the end host, in Proceedings of the 18th conference on USENIX security symposium. 2009, USENIX Association: Montreal, Canada. p. 351-366.
19. Zhou, S. T. a. M., A Heuristic Approach for Detection of Obfuscated Malware,. IEEE, 2009.
20. [20] Ahmed, M., et al. NIDS: A Network Based Approach to Intrusion Detection and Prevention. in Computer Science and Information Technology - Spring Conference, 2009. IACSITSC '09. International Association of. 2009.
21. Garfinkel, T. and M. Rosenblum, A virtual machine introspection based architecture for intrusion detection. 2003: p. 191--206.
22. Lagar-Cavilla, H. A., Flexible Computing with Virtual Machines. 2009.
23. Gorodetsky, V., et al., Multi-agent Peer-to- Peer Intrusion Detection Computer Network Security, V. Gorodetsky, I. Kotenko, and V. A. Skormin, Editors. 2007, Springer Berlin Heidelberg. p. 260-271.
24. Ye, D., An Agent-Based Framework for Distributed Intrusion Detections. 2009.
25. Ou, C. -M. and C. R. Ou, Agent-Based immunity for computer virus: abstraction from dendritic cell algorithm with danger theory, in Proceedings of the 5th international conference on Advances in Grid and Pervasive Computing. 2010, Springer-Verlag: Hualien, Taiwan. p. 670-678.
26. Bijani, S. and D. Robertson, Intrusion detection in open peer-to-peer multi-agent systems, in Proceedings of the 5th international conference on Autonomous infrastructure, management, and security: managing the dynamics of networks and services. 2011, Springer Verlag: Nancy, France. p. 177-180.
27. Dong, H., et al. Research on adaptive distributed intrusion detection system model based on Multi-Agent. in Computer Science and Automation Engineering (CSAE), 2011 IEEE International Conference on. 2011.
28. Ou, C. M., Multiagent-based computer virus detection systems: abstraction from dendritic cell algorithm with danger theory. Springerlink, 2011.
29. Paritosh Das, R. N., A Temporal Logic Based Approach to Multi-Agent Intrusion Detection and Prevention. 2012.
30. Moser, C. Kruegel, and E. Kirda, “Limits of static analysis for malware detection, ” in Proceedings of the 23rd Annual Computer Security Applications Conference (ACSAC ’07), pp. 421–430, December 2007.
31. P. Okane, S. Sezer, and K. McLaughlin, “Obfuscation: the hidden malware, ” IEEE Security & Privacy, vol. 9, no. 5, pp. 41–47, 2011.

32. S. Cesare and Y. Xiang, Software Similarity and Classification, Springer Science & Business Media, 2012.
33. M. Rajagopalan, M. A. Hiltunen, T. Jim, and R. D. Schlichting, "System call monitoring using authenticated system calls, " IEEE Transactions on Dependable and Secure Computing, vol. 3, no. 3, pp. 216–229, 2006.
34. M. Abadi, M. Budiu, U. Erlingsson, and J. Ligatti, "Control- ´flow integrity," in Proceedings of the 12th ACM Conference on Computer and Communications Security, pp. 340–353, November 2005.
35. S. Sathyanarayan, P. Kohli, and B. Bruhadeshwar, "Signature generation and detection of malware families, " in Information Security and Privacy, Springer, Berlin, Germany, 2008.
36. Sami, B. Yadegari, H. Rahimi, N. Peiravian, S. Hashemi, and A. Hamze, "Malware detection based on mining API calls," in Proceedings of the 25th Annual ACM Symposium on Applied Computing (SAC '10), pp. 1020–1025, ACM, March 2010.
37. Y. Ye, D. Wang, T. Li, and D. Ye, "IMDS: intelligent malware detection system," in Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1043–1047, ACM, August 2007.
38. M. Alazab, S. Venkatraman, and P. Watters, "Zero-day malware detection based on supervised learning algorithms of API call signatures," in Proceedings of the 9th Australasian Data Mining Conference (AusDM '11), vol. 121, pp. 171–182, Australian Computer Society, December 2011.
39. R. Tian, M. R. Islam, L. Batten, and S. Versteeg, "Differentiating malware from cleanware using behavioural analysis," in Proceedings of the 5th International Conference on Malicious and Unwanted Software (MALWARE '10), pp. 23–30, Nancy, France, October 2010.
40. M. Shankarapani, K. Kancherla, S. Ramammoorthy, R. Movva, and S. Mukkamala, "Kernel machines for malware classification and similarity analysis, " in Proceedings of the International Joint Conference on Neural Networks (IJCNN '10), pp. 1–6, July 2010.
41. M. K. Shankarapani, S. Ramamoorthy, R. S. Movva, and S. Mukkamala, "Malware detection using assembly and API call sequences," Journal in Computer Virology, vol. 7, no. 2, pp. 107– 119, 2011.
42. F. Ahmed, H. Hameed, M. Z. Shafiq, and M. Farooq, "Using spatiotemporal linformationin API call swithmachinelea rning algorithms for malware detection, " in Proceedings of the 2nd ACM Workshop on Security and Artificial Intelligence, pp. 55– 62, November 2009.
43. Y. Qiao, Y. Yang, J. He, C. Tang, and Z. Liu, "CBM: free, automatic malware analysis framework using API call sequences," in Knowledge Engineering and Management, pp. 225–236, Springer, Berlin, Germany, 2014.
44. Y. Qiao, Y. Yang, L. Ji, andJ. He, "Analyzing malware by abstracting the frequent item sets in API call sequences, " in Proceedings of the 12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom '13), pp. 265–270, July 2013.
45. J. Bergeron, M. Debbabi, J. Desharnais, M. M. Erhioui, Y. Lavoie, and N. Tawbi, "Static detection of malicious code in executable programs, " in Proceedings of the Symposium on Requirements Engineering for Information Security (SREIS '01), 2001.
46. J. -W. Jang, J. Woo, J. Yun, and H. K. Kim, "Mal-netminer: malware classification based on social network analysis of call graph, " in Proceedings of the Companion Publication of the 23rd International Conference on World Wide Web Companion (WWW Companion '14), pp. 731–734, International World Wide Web Conferences Steering Committee, 2014.
47. K. Rieck, P. Trinius, C. Willems, and T. Holz, "Automatic analysis of malware behavior using machine learning, " Journal of Computer Security, vol. 19, no. 4, pp. 639–668, 2011.
48. C. M. Linn, M. Rajagopalan, S. Baker, C. Collberg, S. K. Debray, and J. H. Hartman, "Protecting against unexpected system calls," in Proceedings of the 14th USENIX Security Symposium, pp. 239–254, Baltimore, Md, USA, August 2005.
49. Anderson, B., Quist, D., Neil, J., Storlie, C. and Lane, T. (2011) Graph Based Malware Detection Using Dynamic Analysis. Journal in Computer Virology, 7, 247-258. http://dx. doi. org/10. 1007/s11416-011-0152-x
50. Bayer, U., Comparetti, P. M., Hlauschek, C. and Kruegel, C. (2009) Scalable, Behavior- Based Malware Clustering. Proceedings of the 16th Annual Network and Distributed System Security Symposium.
51. Anubis. http://anubis. iseclab. org/
52. Indyk, P. and Motwani, R. (1998) Approximate Nearest Neighbor: Towards Removing the Curse of Dimensionality. Proceedings of 30th Annual ACM Symposium on Theory of Computing, Dallas, 24-26 May 1998, 604-613.
53. Biley, M., Oberheid, J., Andersen, J., Morley Mao, Z., Jahanian, F. and Nazario, J. (2007) Automated Classification and Analysis of Internet Malware. Proceedings of the 10th International Conference on Recent Advances in Intrusion Detection, 4637, 178-197. http://dx. doi. org/10. 1007/978-3-540-74320-0_10
54. Park, Y., Reeves, D., Mulukutla, V. and Sundaravel, B. (2010) Fast Malware Classification by Automated Behavioral Graph Matching. Proceedings of the 6th

Annual Workshop on Cyber Security and Information Intelligence Research, Article No. 45.
55. Firdausi, I., Lim, C. and Erwin, A. (2010) Analysis of Machine Learning Techniques Used in Behavior Based Malware Detection. Proceedings of 2nd International Conference on Advances in Computing, Control and Telecommunication Technologies (ACT), Jakarta, 2-3 December 2010, 201-203.
56. Nari, S. and Ghorbani, A. (2013) Automated Malware Classification Based on Network Behavior. Proceedings of International Conference on Computing, Networking and Communications (ICNC), San Diego, 28-31 January 2013, 642-647.
57. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. and Witten, I. (2009) The WEKA Data Mining Software: An Update. ACM SIGKDD Explorations Newsletter, 10- 18.
58. Lee, T. and Mody, J. J. (2006) Behavioral Classification. Proceedings of the European Institute for Computer Antivirus Research Conference (EICAR'06).
59. Santos, I., Devesa, J., Brezo, F., Nieves, J. and Bringas, P. G. (2013) OPEM: A Static- Dynamic Approach for Machine Learning Based Malware Detection. Proceedings of International Conference CISIS'12- ICEUTE'12, Special Sessions Advances in Intelligent Systems and Computing, 189, 271- 280.
60. Islam, R., Tian, R., Battenb, L. and Versteeg, S. (2013) Classification of Malware Based on Integrated Static and Dynamic Features. Journal of Network and Computer Application, 36, 646-556. http://dx. doi. org/10. 1016/j. jnca. 2012. 10. 004
61. Anderson, B., Storlie, C. and Lane, T. (2012) Improving Malware Classification: Bridging the Static/Dynamic Gap. Proceedings of 5th ACM Workshop on Security and Artificial Intelligence (AISec), 3-14.

Global Journal of Computer Science and Technology: E
Network, Web & Security
Volume 16 Issue 5 Version 1.0 Year 2016
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Comparative Analysis: Heart Diagnosis Classification using Bp-LVQ Neural Network Models for Analog and Digital Data

By D. Rajeswara Rao & Dr.JVR Murthy
KL University

Abstract- Decades onwards companies are creating massive data warehouses to store the collected resources. Even though the stored resources are available, only few companies have been able to know that the actual value stored in the database. Procedure used to extract those values is known as data mining. We use so-many technologies to apply this data-mining technique, artificial neural network(ANN) also includes in this data-mining techniques ,ANN is the information processing units which are similar to biological nervous systems. Backpropagation is one of the techniques that used for classification and LVQ (learning Vector Quantization) can be plotted under the competitive learning scheme which is also used for classification. This paper elaborates artificial neural networks, its characteristics and working of backpropagation and LVQ algorithms. In this paper we show the intriguing comparisons between backpropagation and LVQ (Learning Vector Quantization) for both analog and digital data. It also attempts to explain the results between back-propagation and LVQ.

Keywords: artificial neural networks (ANN),activation function,multi-layer-feedforward-network,sigmoid, least mean squared error, backpropagation, training, codebook, competitive networks, learning vector quantization.

GJCST-E Classification : F.1.1 C.2.1

COMPARATIVEANALYSISHEARTDIAGNOSISCLASSIFICATIONUSINGBPLVQNEURALNETWORKMODELSFORANALOGANDDIGITALDATA

Strictly as per the compliance and regulations of:

Comparative Analysis: Heart Diagnosis Classification using Bp-LVQ Neural Network Models for Analog and Digital Data

D. Rajeswara Rao α & Dr.JVR Murthy σ

Abstract - Decades onwards companies are creating massive data warehouses to store the collected resources. Even though the stored resources are available, only few companies have been able to know that the actual value stored in the database. Procedure used to extract those values is known as data mining. We use so-many technologies to apply this data-mining technique, artificial neural network(ANN) also includes in this data-mining techniques ,ANN is the information processing units which are similar to biological nervous systems. Backpropagation is one of the techniques that used for classification and LVQ (learning Vector Quantization) can be plotted under the competitive learning scheme which is also used for classification. This paper elaborates artificial neural networks, its characteristics and working of backpropagation and LVQ algorithms. In this paper we show the intriguing comparisons between backpropagation and LVQ (Learning Vector Quantization) for both analog and digital data. It also attempts to explain the results between back-propagation and LVQ.

Keywords: artificial neural networks (ANN),activation function,multi-layer-feedforward-network,sigmoid, least mean squared error, backpropagation, training, codebook, competitive networks, learning vector quantization.

I. Introduction

Artificial neural networks (ANN), is often called as "neural networks", is a data processing model based on the biological neural network modeling[5]. Neural networks are widely pre-owned to understand the patterns and the connections in the data. The data may be the outcome of a market research effort, etc. Artificial neural networks have been successfully solved many complex practical issues. The Small processing units present in the network are called as "Artificial Neuron", which operates the information using a connectionist approach to perform complex computations[1][5]. Basically, neural network have layered architecture with interconnected neurons as from fig-1.1. The neural networks (ANN) can be generally be a either a multiple-layer or a single-layer networks. The multilayer structure of neural networks is shown in fig-1.1.

Author α: Department of Computer Science and Engineering, KL University, Andhra Pradesh. e-mail: rajeshduvvada@kluniversity.in
Author σ: Professor, Dept of computer science & engineering, JNTU-Kakinada, Andhra Pradesh.

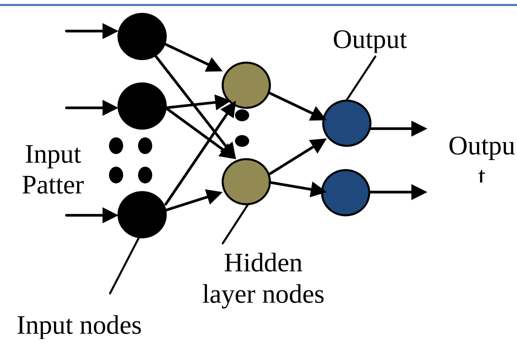


Fig. 1.1 : Architecture of Neural networks

Artificial neural networks had been developed based on the following hypothesis:

- The information is processed among many simple processing units, well known as "neurons".
- The signals are processed among these processing units which are known as neurons over the connection links among them.
- Each and every connection link among these neurons contains an weight, multiples with the transmitted signal.
- Each and every neuron or processing unit applies activation function to its net-input(weight multiplied with its signal input) comes from its previous unit.

Let consider a neuron h1 from fig-1.2, which receives inputs from input neurons y1,y2,y3. The weights on the connection from y1,y2,y3 are w1, w2, w3. The net-input N_y from the input nodes with the activations Y1,Y2,Y3 to the neuron h1 is defined as follows:

$$N\_y = w1Y1 + w2Y2 + w3Y3.$$

As from the final assumption pass this net input to the activation function given as h1= f(n_y).

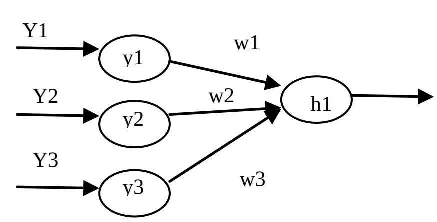


Fig-1.2 : Neuron output generation in ANN

Global Journal of Computer Science and Technology (E) Volume XVI Issue V Version I Year 2016

Some simplifications are necessary to understand the intended properties and to attempt requires mathematical analysis. To implement the above assumptions the whole process of the neural networks are divided in to building blocks. The main building blocks of the neural networks are as follows:

- The Architecture of network.
- Initializing the weights to the nodes.
- Activation Functions.

a) Architecture of Neural Networks

The settlement of the neurons into several layers and the arrangement of the connection within and in-between the layers are known as the network architecture. The basic architecture of the simplest possible neural networks that performs classification subsists of a input layer units and a single output layer unit. Number of layers in the neural network can be outlined as the number of layers, which has weighted interconnected links among the neurons. Advanced neural network architecture consists of hidden layer along with the input layers and output layers. If the two layers of interconnected weights are present, then it is found to have hidden layer. The network architecture is divided into different types like Feed-Forward, Feed-back, Competitive. For back-propagation algorithm we are using Feed-Forward algorithm, where to LVQ (learning Vector Quantization) uses competitive network.

- Feed-Forward networks: These feed-forward networks have either a single layer of weights, where the neurons in the input layer are directly having connection links to the neurons in the output layer, or multiple layers with an interceding set of hidden neurons. Feed-back networks are also associated in two different ways i) Singlelayer ii) Multilayer. As in the single-layer feed-forward networks the weights from input layer does not influence the output layer. Whereas in multilayer feed forward networks one or many layer of nodes (units) between the input layer and the output layer units, so this network is used to solve the complex problems.

b) Setting the weights to the nodes:

The process of setting the weights enables the learning rules or training process. A neural network focusses on the way in which the weights can be changed. The method of tuning the weights on the connections among the network layers to attain the expected output is known as the network training. The internal process in the network training is called as learning. Basically, the training process is divided into three types i) supervised ii) Unsupervised and iii) Reinforcement training. For both Back-propagation and for LVQ we are using supervised learning to train the data.

Supervised Learning Rule: It is a procedure of contributing the networks with a sample of inputs and collating the output with a target output. Training process continues until we get the target output. The weights must adjust according to the algorithm. The various learning rules that follow the supervised learning are Delta rule, generalized delta rule, Competitive learning rule. Generalized delta rule is used to train the given data set in the back propagation algorithm, where as competitive learning is the process used to train dataset used for LVQ.

- Delta-Rule: This rule is purely based on the least mean squared error (LMS). The Mean squared error is nothing but the average of all the errors calculated between the target and actual values. This rule is used to minimize the error. Let discuss in detail, for a taken input data the output data is equated with a target output. If the difference between target and actual data is zero, no learning process is considered, otherwise the values of weights are adjusted to lessen the error obtained. The difference between the target output to the actual output value is defined as Δ (wij) = n* k_i *er_j, where n is the learning rate (α), k_i is the activation of unit and er_j is the difference between the target value and actual output value. This learning rule not only progress the weight vector nearer to the target weight vector, it does so in the most efficient way.

Generalized delta rule: Actually the delta rule uses the local information about the error, where the generalized delta rule deals with error information that is not local. The rule is stated in simple sense as follows for weights updating in a cycle after all the training patterns are presented as $W^{new}=w^{old} - n*(E(k))$ where n is learning rate and E(k) is the error difference between the target and actual output.

Competitive Learning Rule: In this competitive learning rule, the neurons present in the output-layer of the neural network compete among themselves to be in an active-state. The major idea behind this rule is that to allow the processing units (neurons) to challenge for the authority to answer a taken sets of inputs, such that only a output neuron (processing unit) challenge for the right to respond for a given subset of inputs. So that only a neuron in the output-layer is in an active-state at a time. The neuron which wins in the competition is known as winner-takes-all neuron. Let W_{kj} denotes the weight of input-layer node (unit) j to neuron. The neuron learns by altering the values of weights from inactive input mode to active input mode. If a neuron (processing unit) does not give acknowledgement to a particular input layout, then the learning does not happens in that particular neuron. If any of the neuron wins in the competition, then its weights are adjusted as follows.

ΔW_{kj}= n (X_j-W_{kj}), when neuron k wins the competition.

=0 ,when neuron k losses the competition.

As from above formulae "n" is well known as the learning- rate(α). The values of the weights are initially set to random values and those weights are being normalized during learning phase (either supervised or unsupervised). The winner-takes-all neuron is selected by using Euclidean distance.

c) Activation Function

The activation function is used to calculate the output comeback generated by neurons. Threshold function performs final mapping of activations of network neurons. The outcome of any neuron is a result of thresholding (internal activation). The aggregate of the weighted input signals is pertained with activation function to get the response. There may be linear and non-linear activation functions. Generally, the activation functions are classified into different types[2]:

i. Identity Function.
ii. Binary Step-Function.
iii. Bipolar step function.
iv. Sigmoidal function.
v. Ramp function.

We use sigmoidal function for the backpropagation algorithm, competitive activation function for the LVQ.

Sigmoid function: Generally these functions are represented by S-shaped curved. These functions are differentiated by its output ranges. Hyperbolic tangent activation function is the most important of all the sigmoid functions with the range of (-1,1). Logistic function has its range of values in between (0, 1). These functions are represented as follows:

a. *Hyperbolic tangent sigmoid activation function:* $S(y)=tanh(y)= (e^{y}-e^{y})/(e^{y}+e^{y})$
b. *Logistic sigmoid function:* $S(y)=1/(1+e^{yx})$.

The graphical representation of above sigmoid functions is shown in following FIG:1.3 (Logistic Function) , fig: 1.4 (Hyperbolic tangent function)

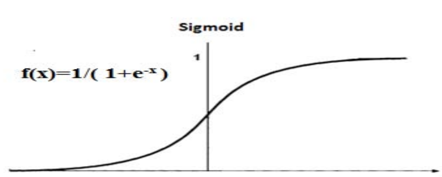


Fig:1.3 : Logistic sigmoid function

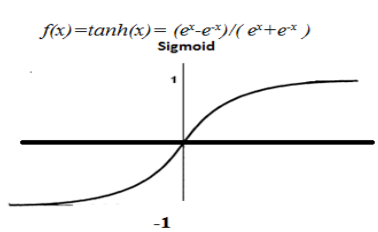


Fig:1.4 : Hyperbolic tangent sigmoid function

The representation of range of the each activation function are defined in table :1.1.

Table 1.1 : Description of activation functions

Function	Definition	Range
Identity	X	(-∞,∞)
Logistic	$S(x)=1/(1+e^{-x})$.	(0,1)
Hyperbolic tangent	$S(x)=(e^{x}-e^{-x})/(e^{x}+e^{-x})$	(-1,1)
Ramp	R(x)= x , x>=0 =0 , x<0 R(x)=max(x,0)	[-1,+1]
Step	O, if x<01, x>=1	[0,+1]

II. Overview of Backpropagation and LVQ

a) Backpropagation Algorithm

Backpropagation is one of the neural network learning algorithms, delineated to diminish the mean square error. Backpropagation is also well-known as the "error backpropagation", because this algorithm is purely based on the error correction learning rule. This algorithm is used to train the multi-layer artificial neural network. Back propagation uses supervised learning rule, in which it generates error by comparing target output to actual output. The backpropagation algorithm could be broken down into four main steps[1][2]:

- Initialization of weights and bias.
- Implementation of feed forward technique to input training patterns.
- The method of calculating and backpropagating the associated errors.
- Weights Updation.

During the first stage, the weights are set-up to some random values (e.g., they ranges from [-1.0,1.0]or[-0.5,0.5])[2].Every processing unit in the network is associated with a bias (threshold), which is used to generate the net input. After the initialization of weights and bias, each training tuple is processed by remaining steps. First of all, the training tuple is pass to the networks input layer. During the process of feed forward of input training patterns, each input unit encounters an input signal and transfer these signals net-input to every hidden units in the network.Later each hidden unit in the network then figure-outs the activation function response. The activation function is known as the output response of the unit (neurons), where in backpropagation we use sigmoidal activation function. Fig: 2.1 show the neuron output generation in hidden and output layer diagrammatic representation. As from the fig: 2.1 the output of neuron is generated by using activation function i.e,

$$f(wp+b)= 1/(1+1+e^{-n}), \text{ where } n=wp+b.$$

Every hidden unit in the network then figure-outs the activation function as shown above and sends its signal to the output unit. The output unit performs the activation function and generates the outcome of the neural network for the given input pattern.

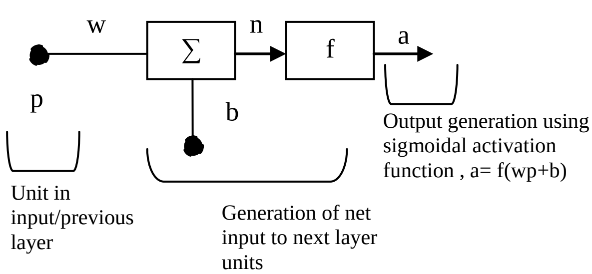


Fig. 2.1 : Activation function generation

During back-propagating the errors, each output units equates its calculated activation function value (a=f(wp+b)) with its target value to determine the error associated with the network. Based on the error, the factor ∂ is computed in backpropagation network for hidden and output layers. As in the final stage , the weights and the bias are updated based up on this factor ∂ and the activation. The backpropagation algorithm implementation is represented in flow chart from fig: 2.2

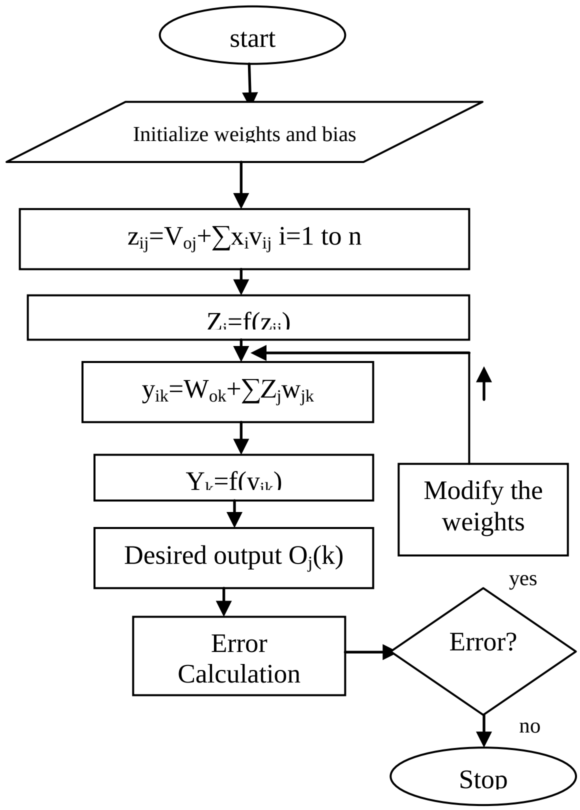


Fig: 2.2 : Flow chart of backpropagation algorithm

The algorithm used in the back-propagation network to train the network is implemented in four different stages is as follows:

Weights Initialization[2] :

Step-1: Initializing the weights and bias to random values (ranges from [-1.0,+1.0] or [-0.5,0.5]).

Step-2: Checking for the stopping condition, if it is false do the steps from 3 to 10.

Step-3: Foe each and every training set, perform the steps from 4 to 9 as mentioned below.

Feed-Forward of input training patterns[3]:

Step-4: Each and every input unit accepts the input x_i and transmits that input signal to hidden layer units.

Step 5: Each hidden unit in the network aggregates its weighted input signals. Activation function to z_{ij} is denoted by Z_j

$$z_{ij} = v_{oj}+\sum x_i V_{ij} \quad .i=1 \text{ to } n$$

$$Z_j= f(z_{ij})$$

The result obtained from this activation function is the input to next layer in the network.

Step 6: Each output unit in the network, aggregates its weighted input signals . Activation function applied to y_{ik} is denoted byY_k

$$y_{ik}= w_{ok}+\sum Z_j W_{jk}$$

$$Y_k=f\ (y_{ik})$$

Backpropagation of the errors:

Step 7: Error is calculated as

$$E(k)= \sum[O_j(k)-T_j(k)]2 \quad j=1 \text{ to } m$$

$$E=E(k)\ f(y_{ik})$$

Step 8: Find the mean squared error

$$E_t=1/2 \sum E \quad k=1 \text{ to } N$$

Updating of weights and bias

Step 9: For the Output layer the weights and the bias are updated as follows

$\Delta W_{jk}=\alpha E_t z_j$. Updated weight is as follows W_{jk} (new) = W_{jk}(old) + ΔW_{jk}

$\Delta wok=\alpha E$. To update bias is w_{ok}(new) $=w_{ok}$(old) $+\Delta w_{ok}$

Similarly the values of weights and the bias are updated in the networks hidden layer is as follows:

$\Delta V_{ij}=\alpha E_t x_i$. The new weight is calculated as V_{ij}(new) $=V_{ij}$(old)$+\Delta V_{ij}$

$\Delta v_{oj}=\alpha E$. Updated bias is v_{oj}(NEW)$=v_{oj}$(OLD)$+\Delta v_{oj}$

Step 10: Check the stopping condition.

Based upon the algorithm stated above the terms are defined as

x_i– Inputs that given to the input units.

v_{oj} – Bias used in the hidden layer units.

V_{ij} – Weights used in hidden layer units.

w_{ok} – Bias used for the outputunits.

W_{jk}– Weights that initialized in output layer.

α– Learning rate.

a) Learning Vector Quantization (LVQ)

Learning Vector Quantization (LVQ) algorithm is the prototype based supervised classification algorithm. It is a particular case of artificial neural network, which implements "winner-take-all" principle[2]. Winner-take-all is the computational principle applied by which neurons in layer compete with each other for activation. The neuron with highest activation stays active while other neurons shut down. LVQ is trained to classify the inputs according to the given targets. Training in LVQ occurs by performing the competition between the neurons. LVQ uses Euclidean distance to perform the competition between neurons. LVQ performs the classification for every target output unit by considering its input pattern i.e, it uses supervised learning technique.

LVQ defines the class boundaries based upon its prototypes. The prototypes are determined during the training procedure using a labeled dataset (the dataset that we take for training).LVQ system is represented by protocols which are defined in future of observed data. The class boundaries are not depends not only on prototypes but also on nearest neighbor rule and winner-takes-it-all. Weight vector for an output unit in a network is known as the "codebook vectors (CV)" or "reference". The architecture of the LVQ algorithm is as shown fig:2.3, fig:2.4 :

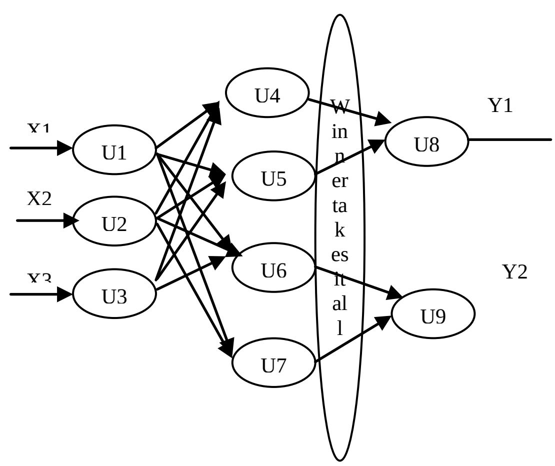


Fig. 2.3 : LVQ architecture

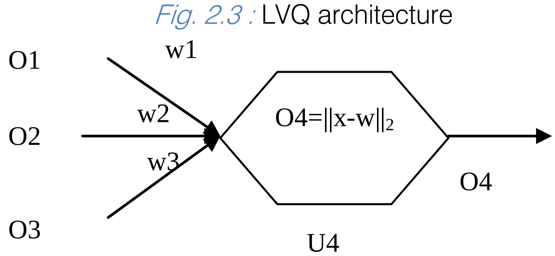
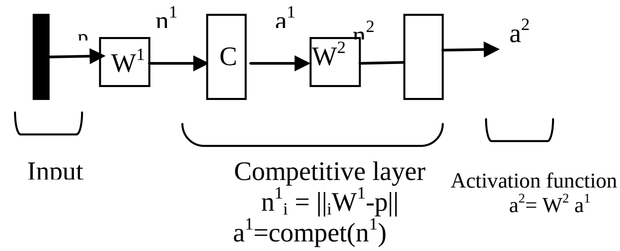


Fig. 2.4 : Inner working of neurons

To express in terms of neural networks, LVQ is a feed-forward neural network. Codebook vector is describe as weight vector(values of weights) of the interconnected weights between all the input layer neurons and hidden neurons. Learning method used in this LVQ algorithm is modifying the weights according to the rules specified and changing the position of code vector (CV) in the input space. Changing the position of CV is nothing but implementing the winner-takes-it-all principle by moving the winner closely if it correctly analyzes the data point or by moving the winner away if it analyzes the data point incorrectly. The working of LVQ is stated diagrammatically in the Fig:2.4



As from the above diagram the net input to the hidden layer is :

$n^1_i = ||_iW^1\text{-}p||$ where $_iW^1$ represents training vector i.e,, inputs given to the input layer p represents Weight vector for the units in next layer it is also called as the codebook vector.

Finally the net output of this input layer is passed to the activation function, where we use the competitive activation function for this LVQ algorithm.

Competitive Activation Function which represents the input/output relation that purely derives by using the Euclidian rule in which

$$a^1 = \text{compet}(n^1)$$

$a^1 =$ 1 neuron which wins the competition

=0 for all neurons.

Therefore the neuron whose weight vector is nearest to the input vector will gives output as 1, and the remaining neurons will gives the output as 0 as shown above. This states that the LVQ network purely competitve network . As initially stated that the neurons in input layer are considered as the same class, after this net output generation to the hidden layer the winning neuron represents a subclass. There may be different neurons that may win the competition, they all belongs to the same sub class.

The hidden layer of the LVQ (learning vector quantization) network combines all subclasses into a single class. As shown in the above figure W^2 done the whole process of combining all the sub classes. W^2 is represented in matrix, in which columns represent the subclasses and the rows represents the classes.

Note: W^2 matrix has a value of 1 in each column, eith the other values set to zero (0).The subclass of a particular class is denoted by the value of 1 in the row. Ex: $W^2_{ij}=1$ means j sub class is a part of ith class.

The input vector **X** is selected at random from the inputs given. If the class labels of the input vector **x** and a codebook vector (weight vector) **W** agree, the codebook vector **W** is moved in the direction of the input vector **X**. If the class labels of the input vector **X** and the codebook vector **w** is disagreed, the codebook vector **W** is moved away from the input vector **X.**

I. Ex: Let $\{W_i\}^1_{i=1}$ stand for the set of weighted vectors (codebook vectors), and the $\{X_i\}^N_{i=1}$ stand for the set of input vectors. Suppose, that the codebook vector W_c is the nearest to the input vector X_i . Let K_{wc} denote the class associated with the codebook vector W_c and K_{xi} denote the class label of the input vector X_i. The values of K_{wc} and K_{xi} are obtained from the W^2 . The codebook vector W_c is regulated as follows:
   If $K_{wc} = K_{xi}$,then $W_c(New) = W_c(Old) + \alpha_n[X_i - W_c(Old)]$ where $0< \alpha_n<1$.
   If $K_{wc} \neq K_{xi}$,then $W_c(New) = W_c(Old) - \alpha_n[X_i - W_c(n)]$,where $0< \alpha_n<1$.
II. Remaining codebook Vectors are not modified.

The learning rate (α) is decreased. This whole LVQ process continues until the stopping condition fails.

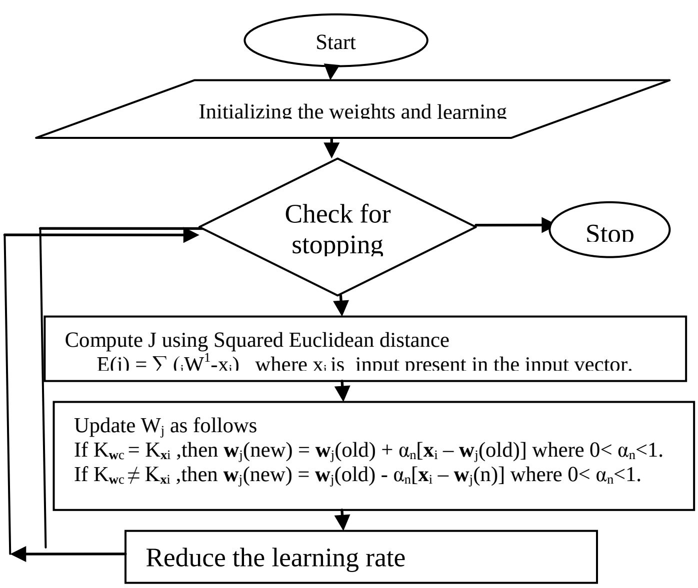
Learning Vector Quantization Algorithm[2]*:*

Step-1: Initialize weights vectors (codebook vectors) and learning rate.

Step-2: Check for the stopping condition. If the condition is false, then perform the steps from 3 to 7.

Step-3: For every training input vector p, do the steps from 4-5

Step 4: Figure out J using Squared Euclidean distance
$E(j) = \sum ({}_jW^1 - X_i)$ where X_i is input present in the input vector.
Find j when E(j)is minimum
Step 5: The value of W_j is updated as follows
If $K_{wc} = K_{xi}$,then $W_j(New) = W_j(Old) + \alpha_n[X_i - W_j(Old)]$ where $0< \alpha_n<1$.
If $K_{wc} \neq K_{xi}$,then $W_j(New) = W_j(Old) - \alpha_n[X_i - W_j(n)]$ where $0< \alpha_n<1$.
Step 6: Reduce the learning rate.
Step 7: Test for the stopping condition.



III. Comparision Between Backpropagation and LVQ

The practical implementation of back-propagation involves factors like choice of network architecture, momentum factor. While implementing these factors backpropagation algorithm associated with few problems like local minima. A local minimum is the problem that occurs frequently, used to change the weights frequently to minimize the error. As in this local minima, in some cases the error might have to rise part of more general fall. If this is the situation the algorithm will struck and the error will not be decreased further. So, for this drawback LVQ gives best results. In this paper we are comparing the efficiencies obtained for testing the heart disease dataset with both backpropagation and LVQ for the two different ranges (-1,1) and (0,1). The following are the results obtained while comparing the both algorithms. The programming is written for 100 instances of a heart diseases dataset from Cleveland with 14 attributes (13 +class attribute).

a) *BackPropagation*

In our paper we practice backpropagation algorithm with different learning rates and finally conclude, how the efficiency changed based upon the value of alpha (learning rate) . To allow fair comparison between backpropagation and LVQ a wide variety of parameter values are tested for each algorithm.

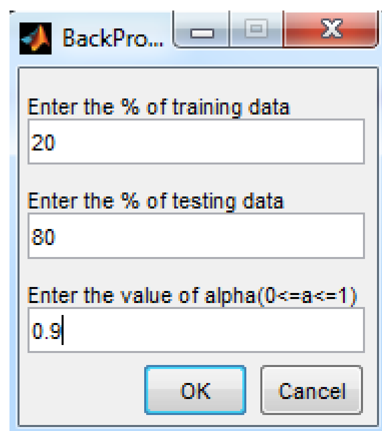


Fig: 3.1 : Input to backpropagation algorithm

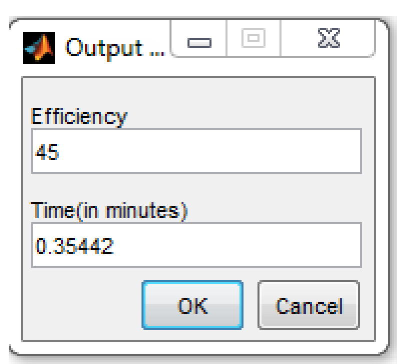


Fig. 3.2 : Output generated for fig:3.1

The backpropagation network is trained on our dataset for different alpha values for different ranges and the observed results are mentioned in the below tables as follows:

When i)α=0.9 (learning rate)

Table. 3.1 : Efficiency obtained for backpropagation (digital) α=0.9

Sl.No	Training(%)	Testing(%)	Time(in minutes)	Efficiency(in%)
1	20	80	2.2	45
2	40	60	0.35	55
3	60	40	0.007	77.5
4	80	20	0.009	75

ii) α=0.8 (learning rate)

Table. 3.2 : Efficiency obtained for backpropagation (digital) α=0.8

Sl.No	Training(%)	Testing(%)	Time(in minutes)	Efficiency (%)
1	20	80	0.003	28.75
2	40	60	0.005	23.333
3	60	40	0.005259	50
4	80	20	0.008362	50

Table. 3.3 : Efficiency obtained for backpropagation (analog) α=0.1

Sl.No	Training(%)	Testing(%)	Time(min)	Efficiency(%)
1	20	80	0.0032099	38.75
2	40	60	0.006441	43.333
3	60	40	0.075057	40
4	80	20	0.010575	60

Table. 3.3 : Efficiency obtained for backpropagation (digital) α=0.1

Sl.No	Training(%)	Testing(%)	Time(min)	Efficiency(%)
1	20	80	0.0032	62.5
2	40	60	0.00503	63.333
3	60	40	0.0066	60
4	80	20	0.00888	79

b) Learning Vector Quantization

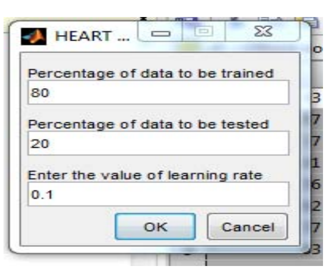


Fig.3.3 : Input to LVQ algorithm

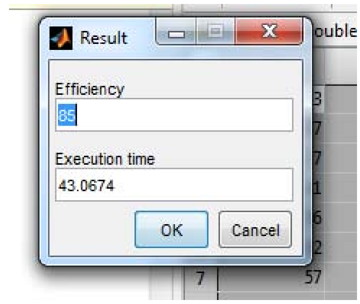


Fig.3.4 : Output generated for fig:3.3

Varying the learning rate alpha from 0.1 to 0.9, it was found that the maximum efficiency is obtained at alpha α=0.1. The results that obtained for various alpha values are shown in the following tables.

Table. 3.4 : Efficiency variations in LVQ analog for α=0.9

Sl.No	Training(%)	Testing(%)	Time(min)	Efficiency (%)
1	20	80	23.7953	54
2	40	60	25.186	57
3	60	40	10.7664	60
4	80	20	10.164	70

Table. 3.5 : Efficiency variations in LVQ analog for α=0.1

Sl.No	Training(%)	Testing(%)	Time(min)	Efficiency
1	20	80	6.5829	64
2	40	60	6.2778	70
3	60	40	8.4187	70
4	80	20	7.175	85

Our paper also attempts to check the efficiency for different ranges i,e for analog (0,1) and bipolar (-1,1). Table:3.3 and Table:3.4 are the results obtained for analog, where the bipolar results are shown in Table:3.5.

Table. 3.6 : Efficiency variation in LVQ bipolar α=0.1

Sl.No	Training(%)	Testing (%)	Time(inmin)	Efficiency(%)
1	20	80	8.7658	70
2	40	60	9.0779	62
3	60	40	12.1897	80
4	80	20	97.8381	70

The better classification efficiency can be achieved by varying the learning rate. As from the above results , we found that the digital gave better efficiency than analog in vector quantization method. It is also found that maximum efficiency was obtained for alpha value 0.1.

IV. Conclusion

In this paper we present a supervised learning based approach to data-mining classification rules for a dataset. The classification is carried out using backpropagation and LVQ. We conclude that LVQ algorithm is one of the best in classification when compared to backpropagation. As from the results obtained for classifying our dataset, we can obtain better classification efficiency by varying the learning rate and it was found that maximum efficiency was obtained for alpha value 0.1 in both algorithms. Comparing the digital results (-1,1) with the analog results, it is found that the digital data gave better efficiency than analog in both back-propagation and LVQ algorithms. Overall comparison between the two algorithms states that the maximum efficiency is obtained in LVQ with high processing time.

References Références Referencias

1. Dr.YashPaul Singh. ,Alok Singh Chauhan (2009)' NEURAL NETWORKS IN DATA MINING',*Journal of Theortical and Applied Information Technology.*
2. S.N.Sivanandam.,S.N.Deepa.,S.Sumathi(2006)*Introduction to Neural Networks Using Matlab 6.0,*Noida: Tata MCGraw-Hill.
3. Athira Mayadevi Somanathan., V.Kalaichelvi(2014) 'An Intelligent Technique for Image Compression', *International Journal for Recent Development in engineering and Technology,*Volume 2(ISSN 2347-6435(online)).
4. A.K.Jain.,J.Mao., and K.M.Mohiuddin,Artificial Neural Networks: A tutorial, IEEE computer VOL.29,no.3,1996,pp.31-44.
5. Priyanka Gaur(n.d)'Neural Networks in Data Mining', *International Journal of Electronic and computer Science Engineering,*IJECSE,Volume 1, Number (ISSN 2277-1956/VIN3-1449-1453),pp. [Online].
6. Agrawal, R., Imielinski, T., Swami, A., "Database Mining : A Performance Persepective", *IEEE Transactions on Knowledge and Data Engineering,* PP.914-925, December 1993.
7. Zurada J.M., "An introduction to artificial neural networks systems", st.paul: West Publishing (1992).
8. Y.Bengio, J.M.Buhmann, M.embrechts, and J.M.Zurada, Introduction to the Special issue on neural networks for data mining and knowledge discovery. *IEEE Trans. Neural Networks.*
9. Haykin, S., *Neural Networks,* Prentice Hall International Inc., 1999.
10. Bradely, I., Introduction to Neural Networks, Multinet Systems Pty Ltd 1997.

Global Journal of Computer Science and Technology: E
Network, Web & Security
Volume 16 Issue 5 Version 1.0 Year 2016
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Identity-based Cryptosystem based on tate Pairing

By Ramesh Ch, K Venugopal Rao & D Vasumathi
GNITS

Abstract- Tate Pairings on Elliptic curve Cryptography are important because they can be used to build efficient Identity- Based Cryptosystems, as well as their implementation essentially determines the efficiency of cryptosystems. In this work, we propose an identity-based encryption based on Tate Pairing on an elliptic curve. The scheme was chosen ciphertext security in the random oracle model assuming a variant of computational problem Diffie-Hellman . This paper provides precise definitions to encryption schemes based on identity, it studies the construction of the underlying ground field, their extension to enhance the finite field arithmetic and presents a technique to accelerate the time feeding in Tate pairing algorithm.

Keywords: identity-based crytosystems, tate pair, elliptic curves and digital certificates.

GJCST-E Classification : E.3 D.4.6

IDENTITYBASEDCRYPTOSYSTEMBASEDONTATEPAIRING

Strictly as per the compliance and regulations of:

Identity-based Cryptosystem based on tate Pairing

Ramesh Ch [α], K Venugopal Rao [σ] & D Vasumathi [ρ]

Abstract - Tate Pairings on Elliptic curve Cryptography are important because they can be used to build efficient Identity-Based Cryptosystems, as well as their implementation essentially determines the efficiency of cryptosystems. In this work, we propose an identity-based encryption based on Tate Pairing on an elliptic curve. The scheme was chosen ciphertext security in the random oracle model assuming a variant of computational problem Diffie-Hellman . This paper provides precise definitions to encryption schemes based on identity, it studies the construction of the underlying ground field, their extension to enhance the finite field arithmetic and presents a technique to accelerate the time feeding in Tate pairing algorithm.

Keywords: identity-based crytosystems, tate pair, elliptic curves and digital certificates.

I. Introduction

The advent of asymmetric encryption represented a great advances in safety of computers, especially because it solved the problem of key exchange algorithms for symmetric encryption. But attacks have been taking the advantage of the fact that it does not have a guarantee on who and the true owner of a public key, so that a user can impersonate another easily by making use of a necessary mechanism of association between a public key and its owner.

To resolve this problem was created the mechanism of certified digital, that uses a hierarchical structure of certifying authorities, able to ensure properly the possession of a given public key. This mechanism works very well in open organizations such as the internet.

In 1984 a model-based cryptographic identities was proposed by Shamir [1]. This model was intended to prevent the use of Digital Certificates, using the identity of the user as its public key. This identity could be an address of e-mail, Social Security number, full name, or a combination is of these elements. The private key would be obtained through a trusted third party(TA - trust authoraty). With this, digital certificates would be necessary only in identification of this central authority, drastically reducing their use. A problem that exists in this idea is the knowledge of the private key by the central authority, needed a total expectations by the user, which requires a lot of care from practical and legal point of view.

On the other hand, does not need the entire infrastructure of hierarchical authorities for the management of the keys by making the model more simple and suitable for organizations where hierarchy and its limitations are well controlled.

Shamir developed a signature scheme based on identities, whose operation is similar to the RSA. He also speculated on the existence of a scheme that has a problem that has been solved in practice by the cryptosystem of Boneh and Franklim [2], whose safety has been rigorously demonstrated.

a) Signature Scheme Based on Identities of Shamir

The signature scheme of Shamir based on Identities and all other forms of encryption based on identities, being divided into four steps:

1. *Setup:* this step and held by authority of expectations to generate the global parameters of the system and the master key, which will underpin that only the TA can generate private keys.
2. *Generation of private key:* this algorithm receives as input the master key and the identity of a user, returning the associated private key.
3. *Signature:* given a private key and a message, the algorithm returns the signature.
4. *Checking:* given an identity, a message and a signature, the algorithm returns true if the signature of that message matches the identity supplied, and returns false if contradicts.

II. Introductory Concepts

a) Security

We will now define some important issues to determine the security of an algorithm based on an additive group, as is the case of elliptic curves encryption [4]:

- *Problem of discrete logarithm:* Given Q = nP, determine n .
- *Problem Computational Diffie-Hellman:* three Data points P, aP, bP, determine abP .
- *Problem of decision Diffie-Hellman:* four Data elements P, aP, bP and cP belonging to a group G, answer true if and only if $C \cong ab(\text{Mod } \#G)$.

One of the first uses of pairings was made by Joux [5]. In this article he showed how the decision has

Author α: Dept. of Computer Science G.Narayanamma Institute of Technology and Science Hyderabad. e-mail: chramesh@gmail.com
Author σ: Dept. of Computer Science G. Narayanamma Institute of Technology and Science Hyderabad.
Author ρ: Dept. of Computer Science Jawaharlal Nehru Technological University Hyderabad.

to be taken to issue the Diffie-Hellman can be easy through the bilinear maps, thus managed to produce an application for key sharing among three parties in a single round.

b) Elliptic Curves

An elliptic curve E defined over a finite field F_p^m and a set of points P = (x, y) with x,y ϵ F_p^m such that $y^2 + a_1 xy + a_3 y + a_2 = x^3 + a^2x^2 + a^4x + a^6$ (standard medium Weierstrass) for $a_i \epsilon$ F_p^m there, beyond the point at infinity, denoted by ∞.

Setting up an operation in an appropriate sum, the elliptic curve form an additive Abelian group with neutral element given by the point at infinity.

An operation widely used in elliptic curve cryptography and scalar multiplication, where a point P and coupled with it own times k to k ϵ Z. A point of order n such that an extent $NP = \infty$ and n the smallest positive integer this property.

III. Identity-based Encryption

The central idea of the public key cryptographic system based on Identity is very simple, because of the fact that the public key is a numeric value without explicit direction and which can be calculated from string of any significance?. In [1], it was proposed that the public key can be the user's identity, such as name , email address , social security number, cell phone number, IP address , serial number of electronic devices, etc.

Is the public key is predetermined (equal to the identity), and then calculate the secret key ? The answer to this question comes with the first model of security assumptions: there is a CA, with the following main responsibilities :

- Create and maintain safe custody of a secret master key S_{AC}
- Identify and record all users of the system
- Calculate the secret keys of the users
- Deliver the secret keys securely (with confidentiality and authenticity)

In 1984, Shamir described the model and algorithms for digital signature. It took almost two decades until efficient encryption algorithms were discovered and demonstrated for the identity -based model to create interest among researchers and industry.

For comparison, in Table 1, we see that the secret key is calculated according to the secret system of authority and the user's identity. For a convenient f, it is not feasible to recover the master key from the ID values . And just the authority is able to generate secret keys, so that secret itself is a guarantee that the use of ID will work in cryptographic operations involving the owners identity.

To encrypt a message to the owner ID or verify a signature ID, user ID using the identity over the public parameters of the system, They include the public key of the authority (see Figure 1).

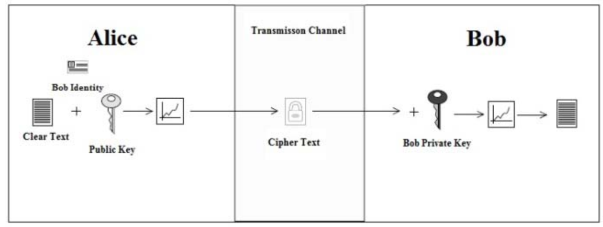
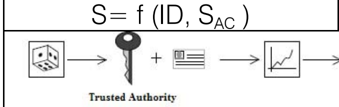


Figure 1 : Encrypting the model based on the identity

Table 1 : Attributes of cryptographic identity -based public key style

Secret key	Public key	Warranty
S= f (ID, S_{AC})	ID	S
		
Calculated by the authority and chosen by the user or shared with the user	Chosen by the user or shared with the user formatted for authority	

To decrypt a message to ID or to create a signature, the secret key ID is required.

a) Advantages

The identity -based model is attractive because it has many interesting advantages. The first is that the public key can in most cases be easily remembered by humans. Very different from the conventional public key, which is usually a binary string with hundreds or thousands of bits? The identity can be informed by the user to their partners and there is no requirement to maintain key directories.

To be able to view the saving processing time, storage costs and data transmissions, we will recall, for example, as It is generally a cryptographic operation with PCI. If Bob wants to encrypt a message to Alice, first of all, he must obtain the certificate that was issued to Alice (consulting a public directory or Alice itself). Bob needs check the validity period and the signature contained in the certificate. The signature verification is a process that sometimes runs the certification path of the certifying authorities involved in the hierarchy until they reach the root certification authority. If nothing goes wrong, Bob can save the Alice certificate for future use. However, before each use, Bob need to consult a validation authority to verify that the certificate has not been revoked (often, a referral to a server that is online). Once the certificate is valid and not revoked, Bob extracts the public key of Alice, encrypts the message and transmits.

In identity -based model, just if the system parameters are authentic Bob can encrypt a message based on the identity of Alice and send (considering that identity withdrawal is treated as explained below).

A peculiarity of identity -based model is that the public key can be used before the secret key calculation. Thus, it is possible to encrypt a message for those who have not registered with the system authority or has secret key for decryption. In contrast to the model based on certificates, the user must first register and get the certificate, and then to receive an encrypted message under your public key.

b) Disadvantages

The first disadvantage, which is characteristic of identity -based systems is the custody of keys. As explained above, the system authority has the ability to generate secret keys of all users under their responsibility. This implies that the authority reaches to the level of confidence that defined in [10]. Consequently, you can decrypt any encrypted texts that have access (if you can identify the recipient's identity). You can also sign on behalf of any user and there is no irreversibility guarantee. Therefore, it is essential that the system of authority is reliable enough for eavesdropping of shares or counterfeiting as these are controllable.

Custody of property keys, referenced by key escrow in English texts is not always undesirable. Within a company, for example, if all sensitive documents and data are encrypted by the employee who created it , the board may have access to decryption in case of death or termination of the employee . When there is need for monitoring the content of encrypted e-mail, it can also be justifiable custody of keys. However, for most applications, custodial key is a disadvantage.

Another point unfavourable to identity -based model is the need for a secure channel for distribution of secret keys. If delivery occurs in networked and remote environment, it is necessary to ensure mutual authentication and delivery with secrecy.

Another concern that one must have in identity -based model is the possibility of identity revocation. If the secret key of a user is compromised, its identity should be repealed. Therefore, it is not recommended to simply use the number of CPF or mobile phone, for example, as a user identifier.

c) Additional features

As noted by [1], the identity -based model is ideal for groups of users, such as executives of a multinational company or branch of a bank, once the headquarters of these corporations can serve as system authority in all trust. Applications small scale, where the cost of deploying and maintaining an ICP are prohibitive, are candidates for the use of identity -based model. When the disadvantages cited above are not critical, the characteristics model allow interesting implementations.

Some examples of services with time availability confidential document that can be revealed to the press or to a particular group , only from certain date and time; bids an auction that should be kept secret until the end of negotiations ; or view a film that should be enabled only within the rental period contracted.

The identity -based model has also been the subject of studies in search for alternatives to SSL / TLS, to Web applications , as shown in [7]. With the elimination of certificates the process of distributing public keys and access control will be simplified. Similarly, the model has been explored to provide security in a number of other application areas , such as grid computing and sensor networks (see for example [5] and [8]) and other applications.

IV. Pairings

A pairing and a pair of mapping linearly independent points of an elliptic curve elements of a finite field is not cyclic. We denote the pairing of two points P and Q e(P, Q). The properties listed below are very interesting for cryptographic applications, are present both in pairing as Weil pairing Tate:

- *Identity:* Pairing a pair of matched points and mapped to the neutral element of the underlying finite field
- *Bilinearidade:* data three points P, Q, R, pairing P + Q and R and the multiplication of the P and R pairing by pairing Q and R. This property is the most important of all, because through it we get the following:

$$e(P,nQ) = e(P,Q)^n = e(nP,Q)$$

- *Do not degeneration:* If P and Q are linearly independent, so their pairing and distinct from the neutral element of the underlying finite field.
- *Efficiency:* data any two points, its pairing can be calculated efficiently by a computer.

a) Tate Pairing

K is an integer such that F_q^k contains the n nth roots of unity. Pairing Tate and defined through the following mapping:

$$e : E[n] \times E/nE \rightarrow F_q^k/(F_q^k)^n$$

where E [n] are the points P of the curve such that nP = ∞.The Tate pairing can be calculated as e(P, Q) = g (D) where D and a divider point Q associated with a function whose rational divider n[P] - n [∞]. The Miller algorithm [Mil04] can be used to calculate the function g.

Menezes, Okamoto, and Vanstone [6] pairings used to perform a transformation of an elliptic curve points supersingular to elements of a finite field generated by the unitary roots of unity. This transformation has allowed a large reduction in the difficulty of the discrete logarithm problem for these curves.

Sakai, Ohgishi and Kasahara [8] made possible the construction of a ciframento protocol based on identities using pairings, this solved the problem proposed by Shamir in his article.

V. Proposed Scheme

Now we can describe in detail the proposed scheme.

Configuration: Given k, the PKG singles groups of bilinear maps, G_1, G_2 and G_t, of prime order $p > 2^k$ generators $Q \in G_2$,$P=Ø(Q) \in G_1$,$g=e(P,Q) \in G_t$ Select s random belonging to Z^*_p a public key of $Q_{pub} = SQ \in G_2$ system summary cryptographic functions H_1, H_2 and H_3.

Generation of key pair: For an identity ID, the private key and $S_{ID} = \frac{1}{H1(ID)+S}$ $Q \in G_2$.

Encryption: Given a message M , the identity of the sender ID_r and the identity of the recipient ID_d, random x is used belonging to Z^*_p to calculate

$r=g^x$,$C=M \oplus H_3(r)$ and $h = H_2(M,r)$.

It is estimated $S=(x + h)\ \varphi\ (S_{ID})$ and $T = x(H_T(ID_r)P + \varphi(Q_{pub})$.

The ciphertext and the triple (c, S, T).

Deciphering and verification: Given the triple (c, S, T) and the identity of the ID_R sender is calculated as

$r = e(T, S_{IDd})$, $M = c \oplus H_3(r)$ and $h = H_2(M, r)$.

Accept message if $r = e(S, H\text{-}1\ (ID_r)Q+Q_{pub})g^{-h}$, in which case the message M and signature (h, S) are returned.

VI. Review

This proposed scheme is interesting because their safety was demonstrated by Barreto semantically, in order to not be subject to attacks that occur when they are used some optimizations of Weil and Tate pairings. Also, please note that the simple junction of the features of this scheme and signature represents a gain of security.

But there is a problem that has not been discussed, which is the abrogation of the private key. This question this open and represents a major problem for the security of any key establishment protocol, because the User can and should change your private key regularly. The problem is in the fact that the private key calculation is deterministic, that is, given the master key sea identity ID, the algorithm always returns the same private key. As the public key and the very identity, the User can not change your identity to obtain a new private key, and needed some other solution. Other asymmetric encryption schemes do not have this problem because the public key is published and revoked with its corresponding private key.

VII. Conclusion

In this work it was possible to see that cryptosystems based on Identities are very interesting and represent an area of research that is growing. However the joint utilization of digital certificates and Identity-Based Protocols can be even more interesting as these two possible solutions to the problem of ensuring association between public key and its owner seem to be complementary.

References Références Referencias

1. Shamir, "Identity-based cryptosystems and signature schemes", Advances in Cryptology - Proceedings of CRYPTO 84, Lecture Notes in Computer Science, 196 (1985), 47–53.
2. D. Boneh and M. Franklin, "Identity-based encryption from the Weil pairing", Advances in Cryptology – CRYPTO 2001, Lecture Notes in Computer Science, 2139 (2001), 213–229. Full version: SIAM Journal on Computing, 32 (2003), 586–615.
3. K. Paterson and G. Price, "A comparison between traditional public key infrastructures and identity-based cryptography", Information Security Technical Report, 8(3) (2003), 57–72.
4. W. Mao. Modern Cryptography - theory and practice. Prentice Hall, 2004.
5. Joux. A one round protocol for tripartite Diffie-Hellman. In W. Bosma, editor, Algorithmic Number Theory, IV-Symposium (ANTS IV), LNCS 1838, pages 385–394. Springer-Verlag,2000.
6. J. Menezes, T. Okamoto, and S. A. Vanstone. Reducing elliptic curve logarithms to a finite field. In IEEE Trans. Info. Theory, number 39, pages 1636–1646, 1983.
7. L. Adleman and M. Huang, "Function field sieve methods for discrete logarithms over finite fields", Information and Computation, 151 (1999), 5–16.
8. R. Sakai, K. Ohgishi, and M. Kasahara. Cryptosystem based on pairing. In Symposium on Cryptography and Information Security, Okinawa, Japan, January 2000.
9. O. Ahmadi, D. Hankerson and A. Menezes, "Software implementation of arithmetic in F3m", International Workshop on Arithmetic of Finite Fields (WAIFI 2007), Lecture Notes in Computer Science 4547 (2007), 85–102.
10. ANSI X9.62, Public Key Cryptography for the Financial Services Industry: The Elliptic Curve Digital Signature Algorithm (ECDSA), American National Standards Institute, 1999.
11. Atkin and F. Morain, "Elliptic curves and primality proving", Mathematics of Computation, 61 (1993), 29–68.
12. R. Balasubramanian and N. Koblitz, "The improbability that an elliptic curve has

subexponential discrete log problem under the Menezes-Okamoto-Vanstone algorithm", Journal of Cryptology, 11 (1998) 141–145.
13. P. Barreto, S. Galbraith, C. ´O h´Eigeartaigh, and M. Scott, "Efficient pairing computation on supersingular abelian varieties", Designs, Codes and Cryptography, 42 (2007), 239–271.
14. P. Barreto, H. Kim, B. Lynn and M. Scott, "Efficient algorithms for pairing-based cryptosystems", Advances in Cryptology – CRYPTO 2002, Lecture Notes in Computer Science, 2442 (2002), 354–368.
15. P. Barreto, B. Lynn and M. Scott, "Efficient implementation of pairing-based cryptosystems", Journal of Cryptology, 17 (2004), 321–334.
16. P. Barreto and M. Naehrig, "Pairing-friendly elliptic curves of prime order", Selected Areas in Cryptography – SAC 2005, Lecture Notes in Computer Science, 3897 (2006), 319–331.
17. den Boer, "Diffie-Hellman is as strong as discrete log for certain primes", Advances in Cryptology – CRYPTO '88, Lecture Notes in Computer Science, 403 (1996), 530–539.
18. Boldyreva, "Efficient threshold signatures, multisignatures and blind signatures based on the gap-Diffie-Hellman-group signature scheme", Public Key Cryptography – PKC 2003, Lecture Notes in Computer Science, 2567 (2003), 31–46.
19. Boneh, X. Boyen and H. Shacham, "Short group signatures", Advances in Cryptology – CRYPTO 2004, Lecture Notes in Computer Science, 3152 (2004), 41–55.
20. Boneh, G. Di Crescenzo, R. Ostrovsky and G. Persiano, "Public key encryption with keyword search", Advances in Cryptology – EUROCRYPT 2004, Lecture Notes in Computer Science, 3027 (2004), 506–522.
21. Boneh and M. Franklin, "Identity-based encryption from the Weil pairing", Advances in Cryptology – CRYPTO 2001, Lecture Notes in Computer Science, 2139 (2001), 213–229. Full version: SIAM Journal on Computing, 32 (2003), 586–615.
22. D. Boneh, C. Gentry, H. Shacham and B. Lynn, "Aggregate and verifiably encrypted signatures from bilinear maps", Advances in Cryptology– EUROCRYPT 2004, Lecture Notes in Computer Science, 2656 (2003), 416–432.
23. D. Boneh, B. Lynn and H. Shacham, "Short signatures from the Weil pairing", Advances in Cryptology – ASIACRYPT 2001, Lecture Notes in Computer Science, 2248 (2001), 514–532. Full version: Journal of Cryptology, 17 (2004), 297–319.
24. M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

Global Journals Inc. (US) Guidelines Handbook 2016

www.GlobalJournals.org

Fellows

FELLOW OF ASSOCIATION OF RESEARCH SOCIETY IN COMPUTING (FARSC)

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

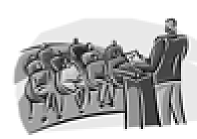
> The “FARSC” is a dignified title which is accorded to a person’s name viz. Dr. John E. Hall, Ph.D., FARSC or William Walldroff, M.S., FARSC.

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

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

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

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

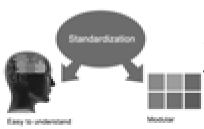


You may join as member of the Editorial Board of Global Journals Incorporation (USA) after successful completion of three years as Fellow and as Peer Reviewer. In addition, it is also desirable that you should organize seminar/symposium/conference at least once.

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

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

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



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

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

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

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

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

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

MEMBER OF ASSOCIATION OF RESEARCH SOCIETY IN COMPUTING (MARSC)

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

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

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

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

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

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

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

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

It is mandatory to read all terms and conditions carefully.

Auxiliary Memberships

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

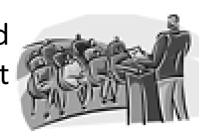
Global Journals Incorporation (USA) is accredited by Open Association of Research Society, U.S.A (OARS) and in turn, affiliates research institutions as "Institutional Fellow of Open Association of Research Society" (IFOARS).

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

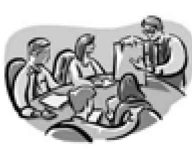
The IFOARS institution is entitled to form a Board comprised of one Chairperson and three to five board members preferably from different streams. The Board will be recognized as "Institutional Board of Open Association of Research Society"-(IBOARS).

The Institute will be entitled to following benefits:

The IBOARS can initially review research papers of their institute and recommend them to publish with respective journal of Global Journals. It can also review the papers of other institutions after obtaining our consent. The second review will be done by peer reviewer of Global Journals Incorporation (USA) The Board is at liberty to appoint a peer reviewer with the approval of chairperson after consulting us.
The author fees of such paper may be waived off up to 40%.



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

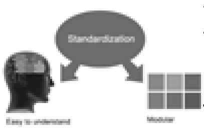


The IBOARS can organize symposium/seminar/conference in their country on behalf of Global Journals Incorporation (USA)-OARS (USA). The terms and conditions can be discussed separately.

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

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

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



After nomination of your institution as "Institutional Fellow" and constantly functioning successfully for one year, we can consider giving recognition to your institute to function as Regional/Zonal office on our behalf.

The board can also take up the additional allied activities for betterment after our consultation.

The following entitlements are applicable to individual Fellows:

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

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

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

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

Other:

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

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

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

Note :

"

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

Process of submission of Research Paper

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

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

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

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

(II) Choose corresponding Journal.

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

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

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

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

Preferred Author Guidelines

MANUSCRIPT STYLE INSTRUCTION **(Must be strictly followed)**

Page Size: 8.27" X 11'"

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

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

1. General,

2. Ethical Guidelines,

3. Submission of Manuscripts,

4. Manuscript’s Category,

5. Structure and Format of Manuscript,

6. After Acceptance.

1. GENERAL

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

Scope

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

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

2. ETHICAL GUIDELINES

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

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

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

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

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

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

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

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

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

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

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

Please mention proper reference and appropriate acknowledgements wherever expected.

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

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

3. SUBMISSION OF MANUSCRIPTS

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

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

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

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

4. MANUSCRIPT'S CATEGORY

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

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

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

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

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

5. STRUCTURE AND FORMAT OF MANUSCRIPT

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

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

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

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

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

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

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

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

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

(h) Brief Acknowledgements.

(i) References in the proper form.

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

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

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

Format

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

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

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

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

Structure

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

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

Abstract, used in Original Papers and Reviews:

Optimizing Abstract for Search Engines

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

Key Words

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

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

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

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

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

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

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

Acknowledgements: Please make these as concise as possible.

References

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

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

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

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

Tables, Figures and Figure Legends

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

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

Preparation of Electronic Figures for Publication

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

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

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

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

6. AFTER ACCEPTANCE

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

6.1 Proof Corrections

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

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

(Free of charge) from the following website:

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

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

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

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

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

6.3 Author Services

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

6.4 Author Material Archive Policy

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

6.5 Offprint and Extra Copies

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

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

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

TECHNIQUES FOR WRITING A GOOD QUALITY RESEARCH PAPER:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Informal Guidelines of Research Paper Writing

Key points to remember:

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

Final Points:

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

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

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

General style:

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

To make a paper clear

· Adhere to recommended page limits

Mistakes to evade

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

In every sections of your document

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

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

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

· Align the primary line of each section

· Present your points in sound order

· Use present tense to report well accepted

· Use past tense to describe specific results

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

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

Title Page:

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

Abstract:

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

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

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

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

Approach:

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

Introduction:

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

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

Approach:

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

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

Procedures (Methods and Materials):

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

Materials:

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

Methods:

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

Approach:

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

What to keep away from

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

Results:

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

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

Content

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

What to stay away from

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

Approach

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

Figures and tables

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

Discussion:

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

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

Approach:

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

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

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

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

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

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

CRITERION FOR GRADING A RESEARCH PAPER (COMPILATION) BY GLOBAL JOURNALS INC. (US)

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

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

Index

save our planet

Global Journal of Computer Science and Technology

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

ISSN 9754350