Volume: 10 Issue: 01 Pages: 3698-3703 (2018) ISSN: 0975-0290

WLI-FCM and Artificial Neural Network Based Cloud Intrusion Detection System

Pinki Sharma
Research scholar, Department of Computer Science, Punjabi University, Patiala, Punjab, India
Email: pinkisharma@gmail.com
Jyotsna Sengupta
Professor, Department of Computer Science, Punjabi University, Patiala, Punjab, India
Email: jyotsna.sengupta@gmail.com
P. K. Suri
Email: pksurikuk@gmail.com

ABSTRACT

Security and Performance aspects of cloud computing are the major issues which have to be tended to in Cloud Computing. Intrusion is one such basic and imperative security problem for Cloud Computing. Consequently, it is essential to create an Intrusion Detection System (IDS) to detect both inside and outside assaults with high detection precision in cloud environment. In this paper, cloud intrusion detection system at hypervisor layer is developed and assesses to detect the depraved activities in cloud computing environment. The cloud intrusion detection system uses a hybrid algorithm which is a fusion of WLI- FCM clustering algorithm and Back propagation artificial Neural Network to improve the detection accuracy of the cloud intrusion detection system. The proposed system is implemented and compared with K-means and classic FCM. The DARPA's KDD cup dataset 1999 is used for simulation. From the detailed performance analysis, it is clear that the proposed system is able to detect the anomalies with high detection accuracy and low false alarm rate.

Keywords - **Cloud Computing, Cloud intrusion detection system, Intrusion Detection System, IDS, Security.**

Date of Submission: April 30, 2018 Date of Acceptance: May 19, 2018

I. INTRODUCTION

In recent years cloud computing has revolutionized the IT world with rapidly emerging and widely accepted paradigm for computing systems. Today numerous organizations have stated to upload their tremendous amount of important data into public cloud. The sensitive information uploaded into public open cloud [1] and that data is vulnerable to many serious security risks such as availability, confidentiality and integrity. The survey By International Data Corporation (IDC)[2] reports that security is the topmost obstacle of cloud computing (Gens, 2009). Furthermore, the continuous uninterrupted service of cloud technology draws the attention of the intruders to obtain entrance and abuse usres assets and services provided by Cloud service provider (CSP). Lockheed martin's (2010)[3] cyber security division white paper shows that major security concern after data security is attack detection and prevention in cloud infrastructure.

Various technologies such as message encryption and firewall protect the network and can be used as first line of defence. Firewall is not suitable for detecting insider attacks. Some of the Denial of Service attacks (DoS) and Distributed Denial of Service attacks (DDoS) are too complex to detect with firewall [4]. Keeping in mind, the end goal to ensure the security of cloud computing environment, it is necessary to develop an intrusion detection system. A traditional network-based or host-based intrusion detection system [1, 5] does not suit virtual cloud environment. In this way, it is imperative to develop an anomaly detection component which is reasonable for detecting the wicked activities in cloud computing systems.

An effective intrusion detection system should be quick, effortlessly configurable, self-checked, hard to cheat, high fault tolerance, accessible without interference, and free from false error with an overhead as least as possible [6]. Its main mean is to assess data frameworks and to perform early identification of noxious action for decreasing the security hazard to an acceptably low level. High false-positive caution rate may trouble data accessibility, though high false-negative alert rate may bring about genuine harm to the secured frameworks as improper access to delicate data and information harming. The performance of IDS is based on the measure of adequate log information, its regular updates on them, and the quick and correct detection of intrusion from the evaluation between current activity of the user and the past data.

In this paper, we design and develop a technique for cloud intrusion detection by means of the WLI fuzzy clustering and neural network. , the WLI fuzzy clustering technique is applied to the cloud computing network to create the distinctive clusters. At that point, the resultant clusters outcome is given as input for the training the neural network for the learning process.

Rest of this paper is organized as follows: Section 2 presents existing approaches to Cloud intrusion detection in cloud. Detailed description of proposed framework is given in section III. Performance and quality results of proposed framework are presented in section IV. Section V concludes the paper with the references at the end.

Volume: 10 Issue: 01 Pages: 3698-3703 (2018) ISSN: 0975-0290

II. RELATED WORKS

N. Pandeeswari and Ganesh Kumar[7] deploy an anomaly detection system called Hypervisor Detector at the virtual machine monitor layer. The Hypervisor Detector is designed with a hybrid approach FCM-ANN which is a combination of Fuzzy C-Means clustering and Artificial Neural Network. This model works in three phases. The first phase of FCM-ANN is fuzzy clustering module which is used to divide the large dataset into small clusters so as to improve the learning capability of ANN. Fuzzy clustering module enhances the performance of artificial neural network. In second phase, various ANN modules are trained according to their cluster values. In third phase, Fuzzy aggregation module is used to combine the results of various ANN. Here, the Hypervisor Detector is compared with Naïve Bayes and classic ANN by using the various evaluation criterions such as precision, recall value and F-value under various attacks. The performance results of FCM-ANN confirm that it outperforms the Naïve Bayes and the classic ANN algorithms even for low frequent attacks. Hence, the proposed Hypervisor Detector is suitable for detecting various attacks with high detection rate and low false alarm rate.

The authors, Vereia et al. [5] have proposed a Grid and Cloud Computing Intrusion Detection System (GCCIDS) that employs an audit system. GCCIDS integrates knowledge and behaviour analysis to discover the intrusions. This system makes use of an event auditor that captures data from various resources like system logs, node messages and services. Based on the captured data, the IDS service can be used to detect intrusions by using behaviour based and knowledge based techniques. GCCIDS uses artificial neural network for behaviour analysis.

Chirag N. Modi et. al.[8] Propose a framework integrating network intrusion detection system (NIDS) in the Cloud. Our NIDS module consists of Snort and signature apriori algorithm. It generates new rules from captured packets. These new rules are appended in the Snort configuration file to improve efficiency of Snort. It aims to detect known attacks and derivative of known attacks in Cloud by monitoring network traffic, while ensuring low false positive rate with reasonable computational cost. We also recommend the positioning of NIDS in Cloud. We present experimental setup and discuss the design goals expected from proposed framework.

Chi-Chun Lo et al [9] proposed the co-operative intrusion detection model for the grid and cloud computing in which the IDS are distributed among the nodes of the grid and alert other nodes when an attack occurs. Indeed, this approach made a giant leap over other models for the same as this helps other nodes in avoiding the same attacks from occurring. This system also helps in preventing single point of failure since the IDSs are distributed across the cloud.

Infan Gul proposed an efficient model that used multithreading technique for improving the performance in the cloud computing environment to handle large number of data packet flows. The researchers have conducted experiments to perform the performance evaluation of their proposed method relative to the single thread approach. They have used parameters like processing time and execution for their comparative study.

Z. Chiba et al. [10] described the Cooperative and hybrid based network IDS system (CH-NIDS) using the Back Propagation Neural network (BPN). They developed the BPN model based on Snort and Optimized method. The snort prior in the BPN was used to detect the unknown attacks. Due to low convergence of BPN, they exploited the optimization algorithm to optimize the parameters which enhanced the detection rate and accuracy. Also, the snort and optimized based BPN was also used to detect the DoS and DDoS attacks by sharing alerts in central log. Thus, simulation results were evaluated to improve the detection rate and mitigate the false rate.

III. PROPOSED HYPERVISOR DETECTOR

The proposed intrusion detection system is developed at the hypervisor layer that uses the proposed model for detecting the intrusion behaviour of the cloud network. The proposed intrusion detection is begin with, the WLI[11] fuzzy clustering technique is apply to the cloud system to produce the distinctive clusters. Then, the resultant clustered result is given as input to the training algorithm for learning process. A back propagation neural network is used for the training purpose.

At first, the input data is provide to the WLI fuzzy clustering method where the data are clustered together to carried out to detect the intrusion. In WLI fuzzy method, the Cluster Validity Index (CVI) is principally used for the clustering of the fed data. Thus, the Euclidean Distance is measured between the data objects, i.e., a pair of centroids or an object centroids are used to evaluate the heterogeneity and homogeneity measures within the clusters. Also, it uses the fuzzy membership function belongs to data object and cluster centroid.

Proposed WLI-ANN

Step 1: Since cluster centroids are randomly generated, the input dataset may not contain the similar clustering results. The N number of clusters are randomly generated from the input data is C_l, $1 \le l \le N$. To enhance the clustering performance, the CVI is used to estimate the index properties of the centroids.

Step 2: The median distance is taken as the principal aspect in the WLI fuzzy clustering method. After that, the distance is measured between the data object and centroid and that is utilised for the separation of different clusters. Accordingly, the fuzzy compactness is resolved with supported by the fuzzy weighting distances [10] and fuzzy cardinality of clusters. The fuzzy weighting distance is measured by,

$$\mu_{ij}^2 \left\| d_i - c_j \right\|$$

where, d_i is the i^{th} data object and c_j represents the j^{th} cluster and μ_{ij} defines the membership function. Then, the fuzzy cardinality of cluster is given as $\sum_{i=1}^{K} \mu_{ij}$.

Volume: 10 Issue: 01 Pages: 3698-3703 (2018) ISSN: 0975-0290

Step 3: Thus, the total fuzzy compactness of the all the clusters ranges from $1 \quad to \quad N$, is expresses as below.

$$WL_f = \sum_{j=1}^{N} \left[\frac{\sum_{i=1}^{K} \mu_{ij} \left\| d_i - c_j \right\|^2}{\sum_{i=1}^{K} \mu_{ij}} \right]$$

Step 4 *:* In order to separate the clusters, the minimum and median distance is measured between the pair of centroids. The distance between N centroids is evaluated by $N(N-1)/2$. The minimum distance of all $N(N-1)/2$ distance is termed as 'min'. Then, the median distance is determined by $\frac{N(N-1)/2}{2}$ distances of all clusters. Thus, the separation measure of the cluster is evaluated as:

$$WL_d = \frac{1}{2}\left(\min_{i \neq j} \left\{ \left\| c_i - c_k \right\|^2 \right\} + median_{i \neq j} \left\{ \left\| c_i - c_k \right\|^2 \right\} \right)$$

Step 5*:* Finally, the WLI fuzzy clustering caters the N number of clusters where the input data are grouped respectively. The WLI is estimated by the ratio of fuzzy compactness and cluster separation. The cluster validity index is determined by,

$$WLI(N) = \frac{WL_f}{2 \times WL_d}$$

The WLI fuzzy clustering mechanism provides the P number of clusters which is then fed into the proposed model. The centroid is selected by the minimum value of WLI value in every cluster. It is formulated by,

$$C = \min_{N \in d_i} \{WLI(N)\}$$

The training algorithm is described below.

- The WLI fuzzy clustering [11] yields the P number of clusters where the input data are grouped together in each cluster. Hence, the ensuing data object is specified as input to the NN model for the training progression. Due to P number of clusters, we require Q number of NN model to train the data.
- In every cluster, the data are grouped in the size of $m_P \times n$, where P defines the total number of clusters. The clustered data is given as input to Q number of clusters. Thus, the clustered output is expressed by,

$$C_j = \{c_{j1}, c_{j2}, \ldots\ldots c_{jk}\}$$

- where, j is the number of output acquired by the WLI fuzzy clustering mechanism and c_{jk} represents the output of k^{th} cluster. Then, the resultant data is fed as input to the proposed NN where the data is trained to detect the malicious activity in the cloud environment.
- Normally, the training algorithm of neural network is mainly used to train the data to perform the classification process.

Once the data are trained in the network, then the trained data are aggregated. The data aggregation is modelled by combining the trained output of Q different NN network models. The intent of data aggregation is to reduce the detection error of the training algorithm. Thus, the aggregated data is fed into the new NN network. The input of new NN is expressed as follows

$$t = \{f_1, f_2, \ldots\ldots, f_q\}$$

where, t is the input of new aggregated model consists of trained data from Q number of NN. Finally, the data size of $m \times 1$ is attained by the aggregation model to perform the intrusion detection. On the other hand, during testing phase, the input data is given into the hypervisor detector where the proposed NN model is significantly detects the intrusions or malicious activity in the cloud network. Based on the above three phases, the intrusion is detected using the Neural network.

IV. EXPERIMENTAL SETUP AND PERFORMANCE

To implement the Hypervisor Detector, this work uses cloud simulator; cloudsim 3.0. The Hypervisor Detector is trained and tested in cloudsim 3.0. To train and test the proposed system, the DARPA's KDD cup dataset 1999[12] is used. This dataset has 41 features and a label specifying the record as either normal or attack.

For testing the system model, the KDD test dataset is used. The performance factors that are frequently used to evaluate the performance of intrusion detection system are as follows.1. True positive rate, 2. True negative rate, 3. False positive rate and 4. False negative rate. True positive rate entails that the intrusion detection system detects true attack that has occurred. True negative rate entails that the detection system has rightly detect the normal condition. False positive rate implies that IDS has mistakenly marked the normal condition as abnormal. False negative rate indicates that the anomaly detection system cannot detect the intrusions after a particular attack has occurred.

ii) Evaluation parameters: The performance of the proposed cloud intrusion detection system is validated by three metrics are accuracy, true positive rate and false positive rate. The description of this metrics is given below.

True Positive Rate (TPR): It is the measure for the extent of positives which are effectively recognized as malignant activity in the cloud environment. It is also termed as sensitivity. The TPR is expressed as:

$$TPR = \frac{TP}{(TP + FN)}$$

False Positive Rate (FPR): It is defined as the probability measures of falsely rejects the normal node in the cloud network. Thus, the FPR is derived by,

Volume: 10 Issue: 01 Pages: 3698-3703 (2018) ISSN: 0975-0290

$$FPR = \frac{FP}{(FP + TN)}$$

Accuracy: The accuracy is the statistical measure of both positive and negative rates. The higher accuracy value provides the better detection performance. It is formulated as given below

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

where, TN and TP are true positive and negatives, FP and FN denotes the false positive and negative value.

The analysed performance is compared with exiting clustering algorithm like FCM[13], KM[14].

V. PERFORMANCE EVALUATION

As shown in figure 1 for the no of cluster 3,4 and 5 K-means attains 93.25, 94.16 and 91.78 TPR , FCM attains 91.16, 90.89 and 87.83 TPR and WLI attains 96.29, 94.42 and 93.56. Compared to K-means and FCM, WLI attains highest TPR. That means WLI outperforms K-means and FCM.

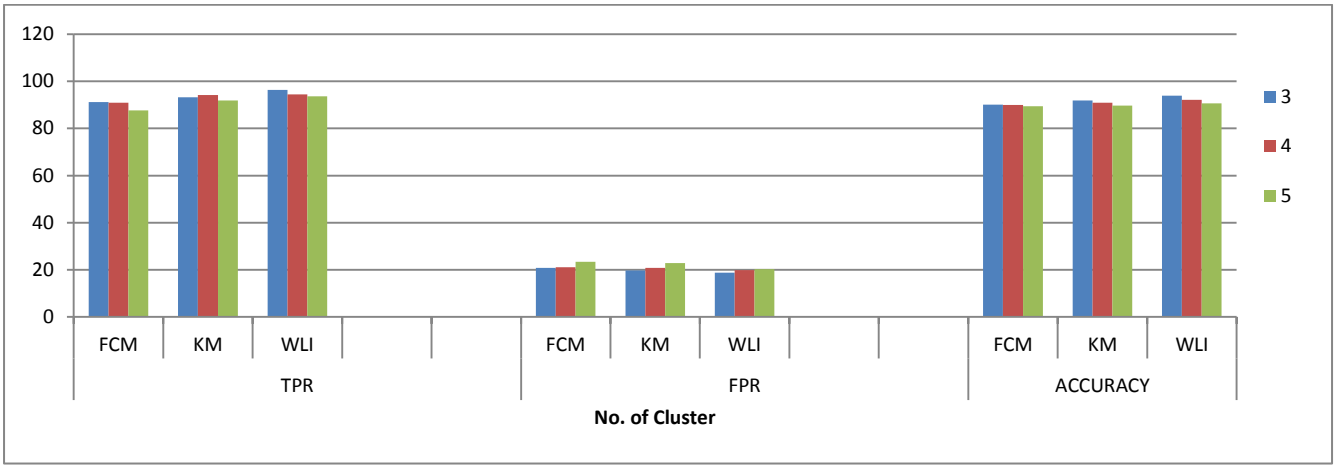

Figure 1. No of Clusters

As shown in figure 1 for the no of cluster 3, 4 and 5 K-means attains 19.8, 20.8 and 22.83 FPR, FCM attains 20.83, 21.13 and 23.41 FPR and WLI attains 18.8, 19.91 and 20.3 FPR. Compared to K-means and FCM, WLI attains highest FPR. That means WLI outperforms K-means and FCM.

As shown in figure 1 for the no of cluster 3, 4 and 5 K-means attains 91.84, 90.94 and 89.67 accuracy, FCM attains 90.15, 89.99 and 89.46 accuracy and WLI attains 93.88, 92.16 and 90.63 Accuracy. Compared to K-means and FCM, WLI attains highest accuracy. That means WLI outperforms K-means and FCM.

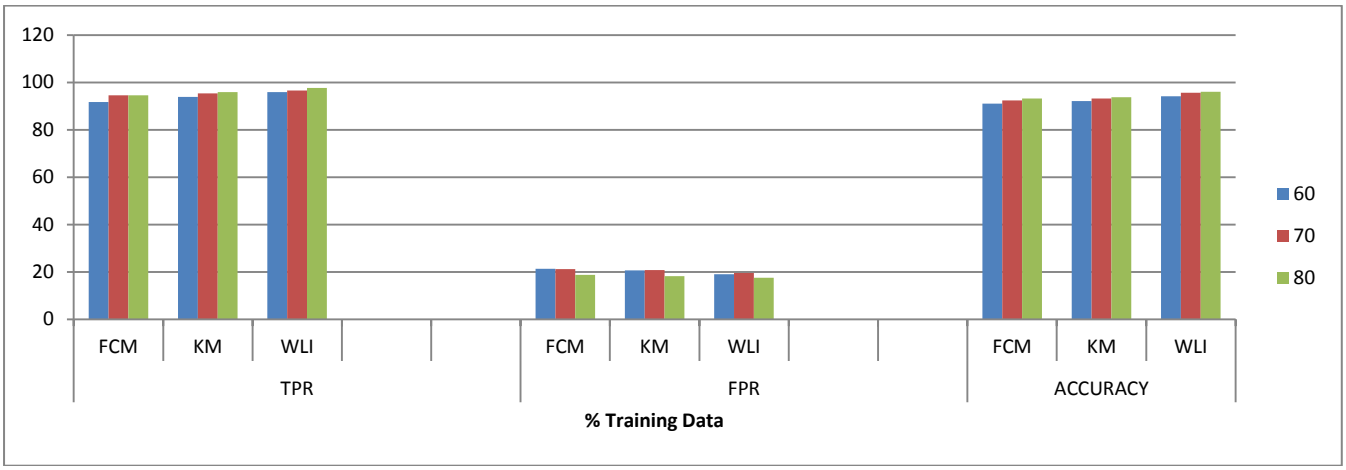

Figure 2. % Training data

As shown in figure 2 for the % Training data 60%, 70% and 80%, the K-means attains 93.86, 95.39 and 95.88 TPR, FCM attains 91.66, 94.51 and 94.56 TPR and WLI attains 95.96, 96.51 and 97.71. Compared to K-means and FCM, WLI attains highest TPR. That means WLI outperforms K-means and FCM.

As shown in figure 2 for the % Training data 60%, 70% and 80, the %K-means attains 20.6, 20.71 and 18.14 FPR, FCM attains 21.3, 21.23 and 18.69 FPR and WLI attains 18.98, 19.5 and 17.46 FPR. Compared to K-means and FCM, WLI attains highest FPR. That means WLI outperforms K-means and FCM.

Volume: 10 Issue: 01 Pages: 3698-3703 (2018) ISSN: 0975-0290

As shown in figure 2 for the % Training data 60%, 70% and 80%K-means attains 92.07, 93.14 and 93.67 accuracy, FCM attains 91.95, 92.37 and 93.26 accuracy and WLI attains 94.18, 95.67 and 96.02 Accuracy. Compared to K-means and FCM, WLI attains highest accuracy. That means outperforms K-means and FCM.

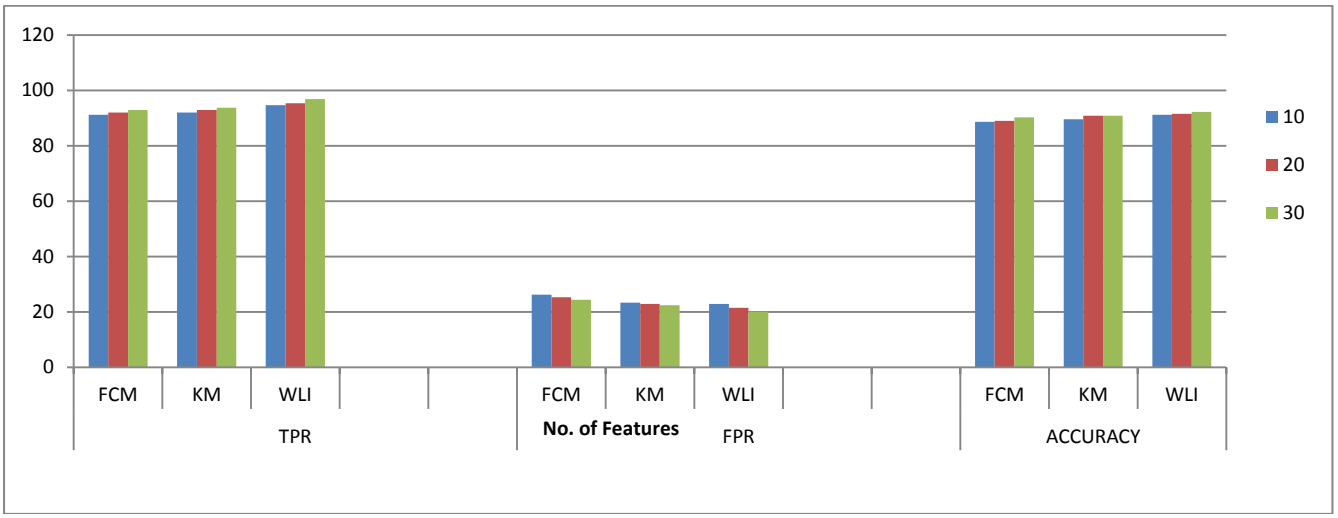

Figure 3 On No. Of features

As shown in figure3 for the Number of Features 10, 20 and 30, the K-means attains 91.94, 92.94 and 93.67 TPR , the FCM attains 91.14, 91.99 and 92.89 TPR and WLI attains 94.67, 95.31 and 96.84. Compared to K-means and FCM, WLI attains highest TPR. That means WLI outperforms K-means and FCM.

As shown in figure 3 for the Number of Features 10, 20 and 30, the K-means attains 23.26, 22.8 and 22.4 FPR, FCM attains 26.16, 25.3 and 24.37 FPR and WLI attains 22.86, 12.51 and 20.12 FPR. Compared to K-means and FCM, WLI attains highest FPR. That means WLI outperforms K-means and FCM.

As shown in figure 3 for the Number of Features 10, 20 and 30, the K-means attains 89.54, 90.87 and 90.8 accuracy, FCM attains 88.67, 88.96 and 90.2 accuracy and WLI attains 91.14, 91.49 and 92.24 Accuracy. Compared to K-means and FCM, WLI attains highest accuracy. That means WLI outperforms K-means and FCM.

VI. CONCLUSION

This paper presents an intrusion detection system called Hypervisor Detector at the hypervisor layer. The Hypervisor Detector is designed with a hybrid approach WLI-ANN which is a combination of WLI and Artificial Neural Network. The fuzzy C mean is running with the WLI. The WLI partially allows the existence of closely allocated centroids in the clustering results by considering not only the minimum but also the median distances between a pair of centroids and therefore possesses the better stability. This model works in three steps. In first step is fuzzy clustering module which is used to divide the large dataset into small clusters so as to improve the learning capability of ANN. In second step, various ANN modules are trained according to their cluster values. In third step, the results of various ANN from the second step are combined to get the final result. The proposed Hypervisor Detector is compared with K-means and classic FCM by using the various evaluation criterions such as number of clusters, number of Features used and % of training data Used. The performance results of proposed WLI-ANN confirm that it outperforms the K-means and the classic FCM algorithms for more TPR, Accuracy and low FPR. Hence, the proposed Hypervisor Detector is suitable for detecting various attacks with high detection rate and low false alarm rate.

REFERENCES

[1] H. Jin, G, Xiang, D. Zou, S. Wu, F. Zhao, M. Li, And W. Zheng, AVMM-based intrusion prevention system in cloud computing environment, *Journal of Supercomputing Springer ,66*(3),2011, 1133–1151.

[2] F. Gens, New IDC IT Cloud Service Survey: Top Benefits and Challenges Exchange,2009,*online; http:// blogs .idc. com/ie/ ?p=730S.(Accessed 12 may 2017).*

[3] L. Martin, WhitePaper,2010, *online:/http://www. Lock heed martin.com /data/assets/isgs/ documents/ Cloud Computing WhitePaper.pdf.*

[4] C. Modi ,D. Patel, B. Borisaniya, H. Patel, A. Patel and M. Rajarajan, A survey of intrusion detection techniques in Cloud, *Journal of Network and Computer Applications, 36*(1),2013, 42-57.

[5] K. Vieira, A. Schulter, C.B. Westphall, and C. M. Westphall, Intrusion detection techniques in grid and cloud computing environment. *IEEE IT Professional Magazine* , 2010,38–43

[6] S.Raja and S. Ramaiah, An Efficient Fuzzy-Based Hybrid System to Cloud Intrusion Detection, International *Journal of Fuzzy Systems, 19*(1),2016,1-16.

Volume: 10 Issue: 01 Pages: 3698-3703 (2018) ISSN: 0975-0290

[7] N. Pandeeswari and Ganesh Kumar, Anomaly Detection System in Cloud Environment Using Fuzzy Clustering Based ANN, *Mobile Networks and Applications, 21*(3), 2016, 494-505.

[8] C. N. Modi, D. R. Patel, A. Patel, and M. Rajarajan , Integrating Signature Apriori based Network Intrusion Detection system (NIDS) in Cloud Computing. *In: Proceedings of 2nd International Conference on Communication, Computing & Security, Procedia Technology,*6:905–912. Doi:10.1016/j. protcy.2012.10 .110

[9] C. C. Lo, C. C. Huang, and J. Ku ,A Cooperative Intrusion Detection System Framework for Cloud Computing Networks, *39th International Conference on Parallel Processing Workshops* , 2010, 280-284.

[10] Z. Chiba, N. Abghour, K. Moussaid and M. Rida, A Cooperative and Hybrid Network Intrusion Detection Framework in Cloud Computing Based on Snory and Optimized back Propagation neural Network, *International Workshop on Mobile Cloud Computing Sytems, Management and Security*, 83, 2016, 1200-1206.

[11] C. Wu, C. Ouyang, L. Chen, and L. Lu, A New Fuzzy Clustering Validity Index with a Median Factor for Centroid-based Clustering, *IEEE Transactions on Fuzzy Systems, 23*(3),2015, 701 – 718.

[12] KDD Cup 1999. Available *online: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html, October 2007.*

[13] R. kulhare and D. Singh, Intrusion Detection System based on Fuzzy C Means Clustering and Probabilistic Neural Network, *International Journal of Computer Applications,74* ,2013, 30-33.

[14] K. Nalavade and B. B. Mehsram, Evaluation of K-Means Clustering for Effective Intrusion Detection and Prevention in Massive Network Traffic Data, International Journal of Computer Applications, 96, 2014, 9-14.

AUTHORS BIOGRAPHY

Pinki Sharma is currently working toward her PhD in the Department of Computer Science at the Punjabi University, India. Her research interests include cloud Computing and information security.

Jyotsna Sengupta is currently working as a Professor in the Department of Computer Science at the Punjabi University, India. She has 30 years of teaching experience with various designations in the Department of Computer Science at the Punjabi University, Patiala and in various other reputed institutes. She received her BE in Electronics and Communications from the Thapar Institute of Engg. And Tech., Patiala, India in 1982 and MS in Computer Science and Engineering from the Santa Clara University, Santa Clara, California, USA, in 1985. She has completed her PhD in Computer Science and Engineering at the Thapar Institute Of Engg. And Tech., Patiala, India, 2002. Her research Interests include cloud computing, ADHOC network, wireless networks and distributed systems. She has Attended a number of national and international Conference and published a number of research paper in National and International journals.

P.K. Suri is a former Professor, Dean Academic and Chairman of the Department of Computer Science and Application, Kurukshetra University and HCTM Technical Campus, India. He has 40 years of teaching and research experience with various designation in the DCSA Kurukshetra University, Kurukshetra and in the HCTM Technical Campus, India. He received his MSc from the IIT Roorkee (formerly known as University of Roorkee), Roorkee, India in 1972. He has completed his PhD at the Faculty of Engineering, Kurukshetra University, Kurukshetra in 1981. His research interests Include simulation, cloud computing, ADHOC network, Wireless networks and distributed systems software engineering. He has attended a number of national and international conference and published a number of research papers in national and international journals. He has guided more than 20 PhD research scholars.