Clustering based Evolving Neural Network Intrusion Detection for MCPS Traffic Security

Nishat I Mowla
Dept. of Computer Science and Engineering
Ewha Womans University
Seoul, Korea
e-mail:
nishat.i.mowla@gmail.com

Inshil Doh
Dept. of Cyber Security
Ewha Womans University
Seoul, Korea
e-mail: isdoh1@ewha.ac.kr

Kijoon Chae*
Dept. of Computer Science and Engineering
Ewha Womans University
Seoul, Korea
e-mail: kjchae@ewha.ac.kr

***Abstract*— In the era of Internet, exploits and vulnerabilities of our systems can be used by attackers to violate confidentiality, integrity, and availability. These attacks pose even more serious consequences when we consider medical networks such as Medical Cyber Physical Systems (MCPS). Therefore, the design of an efficient intrusion detection system is vital. However, the success of most of these systems is linked to custom statistical signature based solutions. It becomes a limiting constraint when there are myriad possible attacks emerging every day. To solve the above issues, several machine learning techniques have been developed to form robust detection systems. Nevertheless, these systems are not efficient with low-frequency attacks and are often considered as outliers, even though the consequences of missing upon such attacks can be dangerous. Therefore, this paper proposes an evolving machine learning technique, based on clustering and neural network classification to improve the detection accuracy of all forms of network intrusion traffic. Our experimental results on the standardized Knowledge Discovery and Data Mining (KDD) Cup 99 public dataset show that the proposed mechanism can outperform the well-established boosted decision tree algorithm under different selected features environments.**

Keywords-Intrusion Detection; Machine Intelligence; Clustering; Neural Networks; Medical Cyber Physical Systems.

I. INTRODUCTION

Network security lies at the heart of the future Internet, as various intrusions in the technological systems can cause fatal damage. Various forms of body worn devices that record multiple physiological signals, such as ECG (Electrocardiogram) and heart rate, or even more sophisticated devices that measure physiological markers such as body temperature, skin resistance, gait, posture, and EMG (Electromyography) are well-connected to the Internet. Medical Cyber Physical Systems (MCPS) combining such sensors aim at providing remote healthcare to patients. Malicious attackers can exploit the vulnerabilities in these networks to breach confidentiality, integrity, and availability. MCPS require assurance of health information privacy during transmission from the sensory network to cloud and from the cloud to the doctor's mobile devices. Therefore, a malicious traffic detection system is vital in such scenarios [1].

The success of most intrusion detection systems is linked to custom signature based solutions. However, it becomes unfeasible when we consider time-critical networks, such as Medical Cyber Physical Systems. Intrusion Detection Systems have been developed over time. They can be divided into two main categories, namely, misuse detection and anomaly detection. Misuse detection systems are based on a signature database of already known attacks. These techniques fail in detecting new forms of attacks. With the emergence of new technologies, such as Cyber Physical Systems and Internet of Things, we are also experiencing new forms of network attacks. On the other hand, anomaly detection works by defining a profile for 'normal behavior' where attacks are detected as deviations from this profile. One of the drawbacks of this technique is that it can incur more false positives and slight deviations of normal instances can affect the detection as they depend greatly on this normal profile [2]. Various data mining approaches have also been proposed over time to detect intrusion. Nonetheless, data combined with machine intelligence has seen a higher success rate. Since networks such as Medical Cyber Physical System can monitor the traffic features over long periods of time, machine learning based intrusion detection systems can form a symbiotic relationship with these networks for creating high performance detection tools.

Following to the stream, we propose a clustering based evolving neural network intrusion detection system leveraging machine intelligence. The idea combines supervised and unsupervised machine learning to work with an evolved pairwise learning approach, which highly enhances the classification borderline. Hence, the technique is used to detect the four major forms of network attacks in different feature selected environments.

We discuss some of the related works in Section II. In Section III, we discuss our proposed mechanism and evaluation results followed by the conclusion in Section IV.

II. RELATED WORKS

A. Intrusion Detection Systems (IDS)

The Intrusion Detection Expert System was first proposed by Dorothy E. Denning in 1986[3]. It was an

 ISBN: 978-1-61208-582-1

expert system to detect known types of intrusions with a statistical anomaly detection component leveraging profiles of users, host systems and the target systems. Subsequently, a new version called Next-Generation Intrusion Detection Expert System was developed [4]. Anomaly detection came into mainstream with DARPA (Defense Advanced Research Projects Agency) Intrusion Detection Evaluation in information security [5]. Later on, it appeared that the DARPA datasets are not appropriate to simulate real network systems. This initiated the need for development of new datasets with a view to developing IDS [6].

B. Machine Learning techniques for IDS

Machine Intelligence has achieved high detection accuracy in developing IDS. The literatures from [7] and [8] discuss a survey of these techniques. One of the most promising techniques among them is the neural network. It consists of a collection of actions to transform a set of inputs to a set of searched outputs through a set of simple processing units, or nodes and connections between them. Both supervised and unsupervised neural network techniques have been developed such as Multi-Layer Perceptron (MLP) [9] and Self-Organizing Maps (SOM) [10] respectively. Neural networks are found to be ideal when we consider all various forms of network attack traffic that we can encounter [11].

Network traffic can sometimes be better represented by clustering techniques where traffic data are clustered and are often unsupervised. There are commonly two main clustering algorithms namely k-means clustering and c-means clustering. Clustering also allows subsampling. Therefore, it can reduce the complexity when fed into a classifier machine. The authors of [12] investigated multiple centroid-based unsupervised clustering algorithms for intrusion detection and proposed a self-labeling heuristic for detecting attacks and normal clusters of network traffic. Clustering techniques are also useful in identifying unseen types of attacks. However, clustering techniques alone are not sufficient to create an effective decision boundary which can achieve promising accuracy rate. Due to these reasons, various hybrid approaches have been developed overtime. The authors of [2] proposed an intrusion detection system using Support Vector Machine and hierarchical clustering where the clustering techniques mainly aided in enhancing the training time of the Support Vector Machine by subsampling of the problem space. Support Vector Machine is an efficient classification technique but it requires higher training time. [13] proposed an intrusion detection technique using ANN (Artificial Neural Network) and fuzzy clustering. In this system, fuzzy clustering technique is used to generate different training subsets which are then trained to formulate different ANN based models. Thereafter, it determines membership grades of these subsets and combines them via a new ANN to get final results. The goal of this mechanism is to increase the detection accuracy of less frequent attacks by evaluating subsets. However, the accuracy of this mechanism increases when the number of clusters is increased, which recurrently incurs computational cost. [14] proposed the use of genetic fuzzy systems and pairwise learning for improving detection rates of low frequency attacks. The pairwise learning approach is used to create m*(m-1)/2 two-class problems for an original m-class problem which is then classified with Genetic Fuzzy Systems (GFS) based on evolutionary algorithm. The pairwise learning approach was helpful to simplify the decision boundary by making the problem space smaller to a two-class problem. Even so, the binarization technique is subject to high computational complexity as the number of total classes will exponentially increase for their proposed two-class problem forming formula.

Neural networks alone perform worse than Support Vector Machine (SVMs), which are outperformed by efficient techniques, such as Decision Tree. Multi-Layer Perceptron (MLP) is one of the simplest Deep Learning Neural Network architectures. In this paper, we have used a fully connected Multi-Layer Perceptron Neural Network with one hidden layer. To reduce the classification complexity provided to the MLP, we have utilized the clustering technique to simplify the decision boundary of our learner tool. Notably, the Clustered Neural Network is applied on an evolved two class problem to leverage the benefits of pair-wise learning approach while the computation complexity of the approach is not subject to increases with an increasing number of class as was identified in [14]. The computational complexity is kept at minimum by maintaining only one two-class problem always. It will be discussed in more detail in the next section. Our proposed mechanism is simple and efficient. It achieves a promising performance in terms of accuracy for all the different attack types including low frequency attacks used in the experiment.

III. PROPOSED MECHANISM

Our proposed mechanism is built on top of a clustering based Neural Network, which essentially clusters an evolved two class problem, which is then trained by a Neural Network model. Therefore, we first discuss our used algorithms before moving on to our proposed model.

A. K-means Clustering and Neural Network

K-means clustering is the widely-adopted technique of clustering input vectors to k number of clusters and can be represented by a summation function as shown in (1),

$$\sum_{i=1}^{n}\sum_{j=1}^{k} u_{ij}^{m} d(\overrightarrow{x_i}, \overrightarrow{c_j}) \quad (1)$$

where n in is the number of objects with k clusters where u_{ij}^{m} is the degree of membership and $d(\overrightarrow{x_i}, \overrightarrow{c_j})$ is the Euclidean distance of vector $\overrightarrow{x_i}$ from cluster centre $\overrightarrow{c_j}$ which can, in turn, be represented as the weighted average of all objects, as shown in (2),

$$c_j = \frac{\sum_{i=1}^{n} u_{ij}^{m} x}{\sum_{i=1}^{n} u_{ij}^{m}} \quad (2)$$

The relationship between u_{ij}^{m} and $d(\overrightarrow{x_i}, \overrightarrow{c_j})$ can be considered as:

$$u_{ij}^{m} \propto \frac{1}{d(\overrightarrow{x_i},\overrightarrow{c_j})} \quad (3)$$

Thus, (3) shows that as the distance between vector $\overrightarrow{x_i}$ and cluster centre $\overrightarrow{c_j}$ increases, the degree of membership u_{ij}^{m} decreases.

On the contrary, Neural Networks classify by training feature inputs through a number of hidden layers to derive higher level features. It can be classified by a non-linear activation function. As shown in Fig 1, x_i are the feature vectors input to the ANN system. In our case, we used 41 features provided by the KDD'99 dataset [15]. KDD'99 is one of the few public datasets that are recognized as standard datasets specifically for intrusion detection [16]. As shown in the figure, u_j and u_k are the hidden layers which are also called the intermediary output layers. u_l is the final output layer which helps us to identify the classes. In this figure, we show two possible output classes by the red and blue circle. w_{ij}, w_{jk} and w_{kl} are the weight from x_i to u_j, u_j to u_k and u_k to u_l respectively which are fine-tuned by Back-propagation algorithm to reduce error in calculating the output.

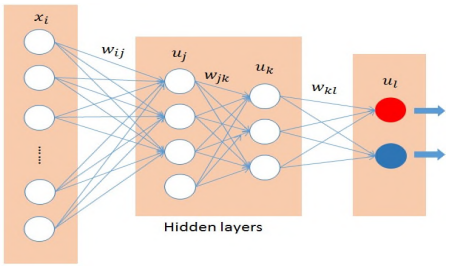

Figure 1. Neural Network classification

B. *Clustering based Evolving Neural Network*

In literature, it has been observed that when a certain classifier is faced with a multi-class problem, it often shows poor results for low-frequency classes. It often happens in case of low-frequency attacks such as U2R (User-to-Root) and R2L (Remote-to-Local) though they are equally fatal to bring a major system down by malicious root access or remote machine access. Hence, we propose an evolving pair of classes to perform a pairwise learning by a Clustered Neural Network. Thus, a single pair of equal size of classes is formed from the standard KDD'99 dataset in order to avoid bias created by low-frequency input vectors. The data is then pre-processed with feature selection. The evolved pairs of classes are then clustered by k-means clustering before classifying them with fully connected neural networks. Fig 2 shows the basic workflow of our evolved pair wise learning with clustered Neural Network.

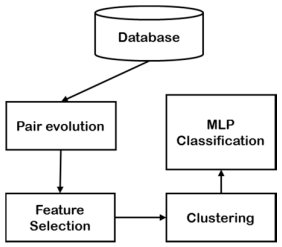

Figure 2. Workflow of pair evolution and training Clustered Neural Network

There are four major types of network attack traffic namely DoS (Denial-of-Service), Probe, R2L and U2R. Among them, DoS refers to all the network traffic flooding attack types. Probe attacks are the attacks conducted by sending meaningless packets in order to gain knowledge about the network. R2L refers to remote access attacks, where the attacker tries to gain access to a remote system. U2R is the type of attack in which the attacker tries to log-in to a normal account and then gain root administrator access [17]. We created our clustered evolving neural network architecture by evolving these four main modalities into a single evolved pair of classes similar to pair evolution algorithm in [18]. Therefore, our first pair of evolved two classes are 'normal' and 'attack'. Here the attack class contains all the four network attacks: DoS, Probe, R2L and U2R. If the tested instance is found not to fall under normal class then the normal class is eliminated from the problem space and a new two class problem is formed from the 'attack' class. Based on prioritization of the attacks, the new two classes are formed. For a certain scenario, let us consider the DoS class to be the most prioritized class. Therefore, the new evolved pair will be 'DoS' and 'other attacks' where the other attacks class contains the other three network attacks: Probe, R2L and U2R. In the next step, if the tested instance is not DoS, we can take the evolved two pair as 'Probe' and 'other attacks' where the other attacks class contains: R2L and U2R. If it is not Probe then we take the network evolved pair as 'R2L' and 'U2R'. In this way, we can make the problem space smaller, which can be better evaluated by our clustered neural network.

IV. Analysis And Simulation

For experimentation, KDD99 dataset with 41 features [17] was used to create a clustered evolving neural network. A total of 10,000 data instances were used.

 ISBN: 978-1-61208-582-1

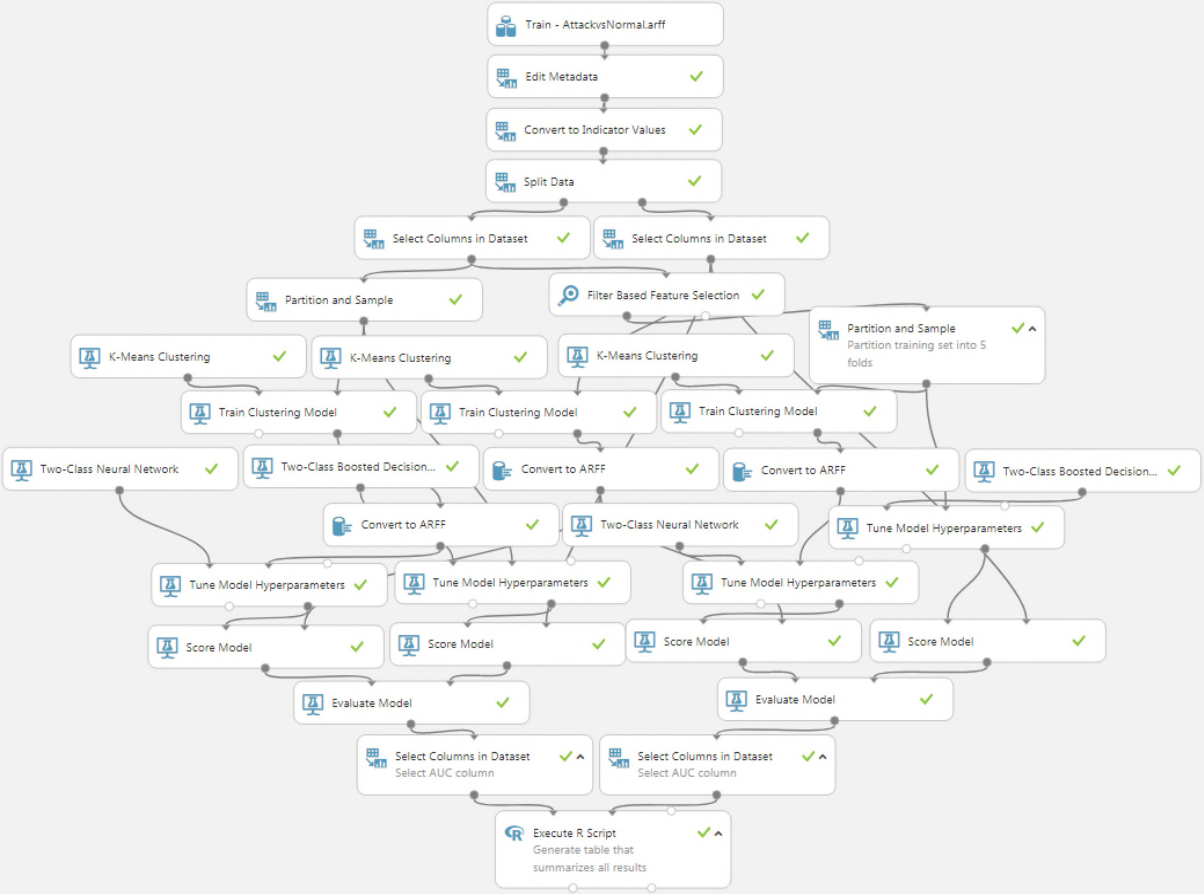

Figure 3. Implementation of Clustered Evoloving Neural Network Intrusion Detection

The data set is split between training and validation set. Therefore, 10% of the data set is used for training and 90% of the dataset is used for validation purpose. Samples from all the subclasses of the 4 major types of network attack traffic were used as shown, in Table 1 [17].

TABLE I. NETWORK ATTACK TRAFFIC

Attack class	Attack Types
DoS	Back, Land, Neptune, Pod, Smurf, Teardrop
Probe	Satan, Ipsweep, Nmap, Portsweep
R2L	Guess_Password, Ftp_write, Imap, Phf, Warezmaster
U2R	Loadmodule

A. Performance Evaluation

We compared our clustered evolving neural network intrusion detection performance with Boosted Decision Tree in two different modes of experiment. In the first experiment, we tested the KDD'99 dataset without feature selection in our proposed environment and in the boosted decision tree environment. In the second experiment, we performed a feature selection method on the dataset to leave it with less number of features. We again compared our proposed model to boosted decision tree. Fig. 4, Fig. 5, Fig. 6 and Fig. 7 show the performance gain in terms of accuracy with clustered neural network in multiple filter based feature selection with Pearson's correlation, i.e., 5 features selection, 10 features selection, 20 features selection and 30 features selection and all features selection. The performances are shown according to the four cases, normal vs attack, DoS vs other attacks, Probe vs other attacks, R2L vs U2R respectively.

 ISBN: 978-1-61208-582-1

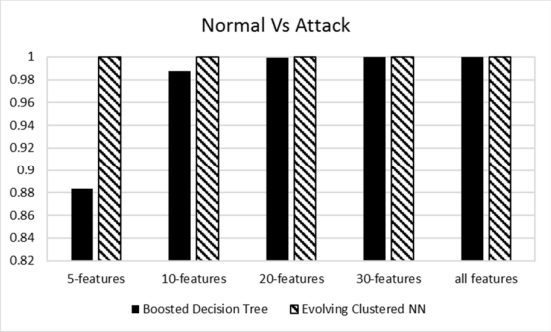

Figure 4. Comparison between Clustered NN and Boosted Decision Tree for Normal Versus Attack

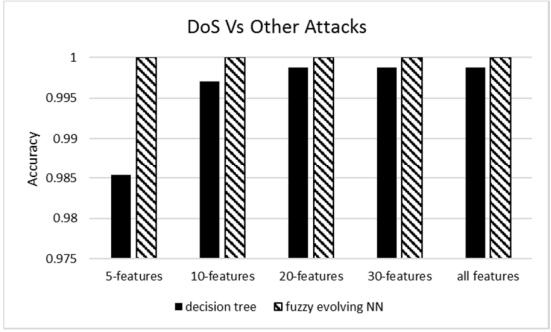

Figure 5. Comparison between Clustered NN and Boosted Decision Tree for DoS Versus Other Attacks

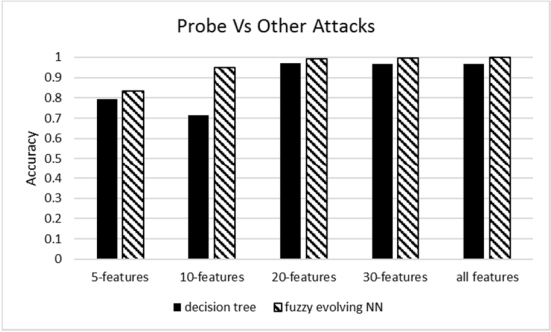

Figure 6. Comparison between Clustered NN and Boosted Decision Tree for Probe Vs Other Attacks

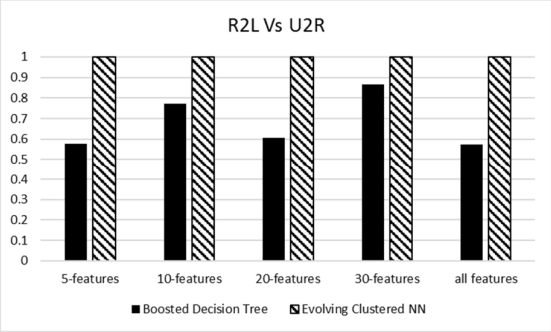

Figure 7. Comparison between Clustered NN and Boosted Decision Tree for R2L Vs U2R

As it can be seen from the above figures, our model has a higher correct classification in all test cases. In Fig. 6, Probe versus other attacks with 5 and 10 features in clustered neural network was found to have slightly lower performance, which we believe could be due to reduced feature size, an essential factor for separation of certain attack categories. However, our proposed mechanism has a higher correct classification when compared to Decision Tree in all the separate experiments of the 4 cases of network attack traffic analysis. This also depicts that if we are limited with the number of features our proposed model may outperform well-known architectures such as Decision Tree. There were no false positives in all the experiments for our proposed model except in the case of Probe Vs Other Attack in 5 and 10 feature cases with 46 cases and 45 cases respectively. The performance gain was exceptionally high for most of our experiment, which could be due to the smaller size of our dataset. Initially, we used 10,000 data instances and it was subject to reduction based on elimination of classes that were not considered to belong to our test instances. To minimize the effect of the size of the dataset, we used 10% of the dataset for training and about 90% of the dataset for validation. Therefore, if we use 1000 instances for training, we used 9000 instances for testing in order to validate the classification methodology in a more constrained environment. Besides, we also tested in different feature selected environments and as can be seen in all cases the performance of our methodology is higher than Decision Tree.

V. DISCUSSION

The performance gain of the method described in this paper is credited to the fact that we decrease the number of concerned classes, thus making the classification simpler. Accordingly, the classifier's complexity is reduced which can be evolved every time to create a two-class problem and solved pairwise to find the specific class of interest. The reduction in complexity is also contributing to the time efficiency of our mechanism. Besides, the elimination process to create a new two-class problem allows us to make the problem space smaller and thus to save more space.

The paper also embraces the idea of combining unsupervised learning with supervised learning by unsupervised clustering of the data before feeding it to the supervised neural network. The prior clustering technique works by creating two subsets where one class is the pure class of concern and the other class is the other class combination. This clustering aids the decision process in neural network by enhancing the classification borderline further and thus achieving higher accuracy.

Finally, the combination of evolved pairwise learning with clustered neural network creates an ultimate leap of performance while reducing the complexity. In this way, it makes the problem space simpler and smaller. The idea, thus, achieves a unique combination of high performance, speed with less space consumption.

Our proposed mechanism, however, does not have any standardized method to prioritize the attack classes which will be given to the evolved two-class pair. Therefore, in future work, we will consider dynamic techniques to

 ISBN: 978-1-61208-582-1

prioritize network attack classes for different network scenarios. We will also consider other emerging attack classes and evaluate our proposed mechanism in such scenarios. Correspondingly, as it was discussed in the performance evaluation section, we will consider bigger initial data instance size for both testing and training for validating our proposed mechanism.

VI. CONCLUSION

In this paper, we have proposed an Intrusion Detection System inspired by evolving a clustered neural network classification technique in order to detect the four key categories of attack traffic that can occur in a Medical Cyber Physical System network. We have presented an enhanced version of the traditional supervised Multi-Layer Perceptron Neural Network developed further when combined with unsupervised clustering. The performance gain has been compared with Boosted Decision Tree in different feature selected environments. To the best of our knowledge, this is the first work done on developing an intelligent intrusion detection system combining evolving pairwise learning with supervised and unsupervised machine intelligence for the Medical Cyber Physical System.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government(MSIP) (No. 2016R1A2B4015899). Kijoon Chae is the corresponding author.

REFERENCES

[1] O. Kocabas, T. Soyata, and M. K. Aktas, “Emerging Security Mechanisms for Medical Cyber Physical Systems”, IEEE/ACM Transactions on Computational Biology and Bioinformatics, Vo. 13, No. 3, June 2016.

[2] L. Khan, M. Awad, and B. Thuraisingham, “A new intrusion detection system using support vector machines and hierarchical clustering” The VLDB Journal—The International Journal on Very Large Data Bases, Vol. 16, No. 4, pp. 507-521, October 2007.

[3] D. E. Denning, “An Intrusion-Detection Model,” in IEEE Symposium on Security and Privacy, pp. 118–131, February 1986.

[4] D. Anderson, T. Frivold, and A. Valdes, “Next generation Intrusion Detection Expert System (NIDES): A summary,” SRI Int., pp. 47, May 1995.

[5] M. Lincoln Laboratory, “DARPA Intrusion Detection Data Sets.” [Online]. Available: https://www.ll.mit.edu/ideval/data/. [Accessed: 07- Apr-2016].

[6] J. McHugh, “Testing Intrusion detection systems: a critique of the 1998 and 1999 DARPA intrusion detection system evaluations as performed by Lincoln Laboratory,” ACM Trans. Inf. Syst. Security., Vol. 3, No. 4, pp. 262–294, November 2000.

[7] J. Singh and M. J. Nene, “A Survey on Machine Learning Techniques for Intrusion Detection Systems,” International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, No. 11, pp. 4349–4355, November 2013.

[8] S. K. Wagh, “Survey on Intrusion Detection System using Machine Learning Techniques,” International Journal of Computer Applications, Vol. 78, No. 16, pp. 30–37, September 2013.

[9] C. Qiu, J. Shan, B. Polytechnic, and B. Shandong, “Research on Intrusion Detection Algorithm Based on BP Neural Network,” International Journal of Security and Its Applications Vol. 9, No. 4, pp. 247–258, 2015.

[10] L. Vokorokos, A. Baláž, and M. Chovanec, “Intrusion detection system using self-organizing map,” Informatica, Vol. 6, No. 1, pp. 1–6, 2006.

[11] J.-P. Planquart, “Application of Neural Networks to Intrusion Detection,” Sans Institute, 2001.

[12] S. Zhong, T. M. Khoshgoftaar, and N. Seliya, “Clustering-based network intrusion detection”, International Journal of reliability, Quality and safety Engineering, Vol. 14 No. 02, pp. 169-187, April 2007.

[13] G. Wang, J. Hao, J. Ma, and L. Huang, “A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering”, Expert systems with applications, Vol. 37, No. 9, pp-6225-6232, September 2010.

[14] S. Elhag, A. Fernández, A. Bawakid, S. Alshomrani, and F. Herrera, “On the combination of genetic fuzzy systems and pairwise learning for improving detection rates on Intrusion Detection Systems”, Expert Systems with Applications, Vol. 42, No. 1, pp. 193-202, August 2015.

[15] TunedIT, “KDD Cup 1999 dataset” [Online]. Available: http://tunedit.org/repo/KDD_Cup/KDDCup99.arff [Accessed: 01- January-2017].

[16] C. F. Tsai, Y. F. Hsu, C. Y. Lin, and W. Y. Lin, “Intrusion detection by machine learning: A review”, Expert Systems with Applications, Vol. 36 No. 10, pp. 11994-12000, 2009.

[17] S. Potluri, C. Diedrich, “Accelerated deep neural networks for enhanced Intrusion Detection System”, 2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA), pp. 1-8, September 2016

[18] N. Mowla, I. Doh, and K. Chae, "Evolving neural network intrusion detection system for MCPS," IEEE International Conference on. Advanced Communication Technology (ICACT), pp. 183-187, February 2017.

 ISBN: 978-1-61208-582-1