International Journal of Research in Advent Technology, Vol.7, No.2, February 2019
E-ISSN: 2321-9637
Available online at www.ijrat.org

A Comprehensive survey of Machine Learning for Intrusion Detection

K. Narayana Rao[1](Research Scholar), Prof. K. Venkata Rao[2], Prof. Prasad Reddy P.V.G.D[3]
Department of Computer Science & Systems Engineering
Andhra University College of Engineering (A)
Andhra University, Visakhapatnam

Abstract: With the massive usage of the internet, the vulnerability of network security becomes an important issue. Intrusion detection System (IDS) is one of the major research problems in network security. IDSs are developed to detect known and unknown attacks of both Computer and computer networks. This paper presents different methods used in IDS for protecting computers and networks for over a decade. This study analyzes different machine learning methods in IDS. It also reviews related studies in the period between 2008 and 2017 focusing on single, hybrid, and ensemble classifiers with relevant datasets.
Keywords: Intrusion detection, Machine Learning, Single Classifiers, Hybrid Classifiers, Ensemble Classifiers

1. INTRODUCTION

Now-a-days internet has become an essential part of our everyday life. It helps people in different areas such as education, business, entertainment etc. In daily life people use the internet applications such as website and e-mail on various activities. As a rapid advancement in internet technology not only giving ease of access to the people, but also sophisticated techniques to the cybercriminals. This leads to the huge number of cyber-attacks on both individuals and organizations.

There are various systems designed to block the internet-based attacks under the internet environment such as firewalls or authentication mechanisms. These techniques are providing some level of security, but they cannot provide protection against inside attacks and malicious code. Particularly Intrusion Detection System (IDSs) helps to detect both internal and external attacks in systems and networks. Intrusion detection is based on assumption that the behavior of intruders who are different from a legitimate user [1].

In general, IDSs can be categorized into two types: misuse (signature) and anomaly detection based on their detection approaches [2] [3]. Misuse detection uses the well-known patterns of attacks to identify the intrusions. On the other hand, anomaly detection tries to identify the activity whether it deviates from the established normal usage pattern.

Based on the literature study the number of intrusion detection systems are developed using different machine learning techniques. Some studies apply single learning techniques as stand-alone, some systems are combining two or more different learning techniques, such as hybrid techniques, and combining multiple weak learners to improve the performance of a classifier known as ensemble classifier. Particularly these techniques are developed as classifiers and clusters. The classifiers techniques are supervised learning that classifies or recognize whether the internet activity is normal or attack, and cluster techniques are unsupervised learning that is trying to find a concealed structure of unlabeled cluster data.

Therefore the aim of this paper is to review related studies published in the past decade by examining the techniques that have been used. Experiments have been conducted, based on the machine learning algorithm perspective what should be considered for future work.

2. MACHINE LEARNING TECHNIQUES

Machine learning is a branch of artificial intelligence that provides to systems, ability to learn automatically and improve from experience without being explicitly programmed. These techniques are used to recognize the pattern. Pattern recognition is a task to take raw data and activity on data category recognition. Supervised and unsupervised algorithms are used to solve different pattern recognition problems. Supervised learning is a task inferring a function from

training labeled data, in which each training data contains a pair of an input vector and class label. Unsupervised learning is used to draw inferences from input dataset without consisting class label. The created model classifies unknown examples into learned class labels.

2.1 Single Classifiers

IDS model develop using one single machine learning algorithm. From literature study, the following different machine learning algorithms are used to solve the problems.

2.1.1 K-nearest neighbor

K-nearest neighbor (K-NN) is a straight forward and traditional non-parametric approach for classification that classifies the objects which are presented as points defined in feature space [4] [5]. From the input vector, it calculates the approximate distances between different points and assigns the unlabeled points to the most frequent class labels among considered in training samples of its k-nearest neighbors. In the process of k-NN model creation, k is an important factor and for different k values generate different performances. If k is large, the model takes large classification time and influence the prediction accuracy. K-NN model is mentioned as an instant based learner not an inductive based learning approach [6]. The k-NN model does not contain the training stage, but only searches input vector and classifies new attributes.

2.1.2 Support vector machine

Vapnik was introduced Support vector machine in the mid-1990. First, it maps the input vector into a higher dimensional feature space and then obtains the optimal separating hyper-plane in the higher dimensional feature space. Rather than whole training samples separating hyper-plane decided by support vector and it is extremely robust to outliers. In general an SVM designed for binary classification that is the training vectors separate into two classes. The support vectors of training samples close to a decision boundary. The SVM provides a penalty factor, it is user-specified parameter and allows to make a tradeoff between the width of decision boundary and number of misclassified samples.

2.1.3 Decision trees

Decision tree creates a classifier through a sequence of decisions, based on already several known instances, in which the current decision helps to make succeed decision. The sequence of decisions are represented in a tree structure. The classification process starts from a root node and ends at a suitable leaf node. Every leaf node has an appropriate classification category. The sample attributes are assigned to each node, each branch value is corresponding to the attributes [6]. It is more acceptable as a single classifier because of its simple implementation. Decision tree is well known as a classification and regression tree i.e tree classifies with a range of symbolic class labels is classification tree and range of numerical values is regression tree [8].

2.1.4 Artificial neural network

An artificial neural network is a unit of processing information. It mimics the neurons of the human brain. For pattern recognition problem multilayer perceptron (MLP) is the most widely used structure of the neural network [9]. The architecture of MLP consists of an input layer, one or more hidden layers with computational nodes, and an output layer. The interconnection between nodes has scalar weights and bias which are adjusted during the training phase. MLP usually train besides with back propagation learning algorithms known as back propagation neural network. Assigns random weights at beginning of training, then adjust weights during train and minimizing the error of misclassification.

2.1.5 Self-organizing maps

A Self - Organizing Map (SOM) is trained using unsupervised competitive learning technique [10]. The aim of the SOM algorithm is to map a high dimension data into two dimension visualization. A Kohonen network is one of the basic type of SOMs, has a feed-forward structure with an input layer and computational Kohonen layer which is designed as a two-dimensional arrangement of neurons. In this structure, each neuron associates to all nodes of the input layer. The network finds the closest node to each training case and moves the winning node which is closest neuron (i.e minimum distance neuron) to the training case. SOM maps similar input patterns are mapped into the same or similar output units. It can self categorize all the inputs providing straight forward methods for data clustering.

2.1.6 Genetic Algorithms

A genetic algorithm is based on evolution and natural selection principles. This concept arrives from the "adaptive survival in natural organisms" [11]. The algorithm starts with randomly generating a huge population of candidate programs. Some type of fitness criteria uses to evaluate the performance of each individual population. The Number of iterations is performed for replacing of low performing programs with genetic recombination of high performing programs. That is low fitness measure programs are deleted because they do not survive for the next computer iteration.

2.1.7 Naïve Bayes

It is simple and commonly used probabilistic classifiers based on Bayes theorem. On the basis of the class label given Naive Bayes assumes that the attributes are conditionally independent and thus tries to estimate the class-conditional probability. The typical structure of naïve bayes represented by a directed acyclic graph (DAG), where system variables are represented by node and influence of one node on another by link [12].

2.1.8 Fuzzy logic

Fuzzy Logic is based on the concept of fuzzy set theory, to occur frequently in the real world. Its range of values lies between 0 and 1. It is effective and very potential technique. It works with human decision-making and reasoning and also uses if then else rules [9] .e.g, the natural event of the raining can be varied from slight to violent. Fuzzy logic has been employed to handle the idea of incomplete truth, where the value of the truth may range between completely true and completely false.

2.2 Hybrid Classifiers

The aim of the IDS is to improve the best possible accuracy for the task at hand. Normally to achieve this objective the researchers design the hybrid classifier problem to be solved. The idea of the hybrid classifiers is to combine two or more machine learning algorithms so that the system improves the performance significantly. In general, the hybrid method consists of two functional components, the first one takes input raw data and produces intermediate results, then the second one takes intermediate results as an input and produces final results [14].

In general hybrid classifiers based on clustering, pre process the data for removing the irrelevant and inconsistent data from training samples of each class. Then the clustered data used as a training sample of classifiers. Finally, hybrid approach is an integration of two or more different techniques in which initially optimizing the performance and finally use model for prediction.

2.3 Ensemble Classifiers

Ensemble classifiers were proposed to improve the performance of single classifiers. It is a process of combining multiple weak learners, trained on different training samples and combining their outputs into a single prediction. Generally, used ensemble methods are bagging, boosting and stacking. Though it is known that the weaknesses of the component classifiers get accumulated in the ensemble classifier, it has been providing an efficient performance in some combination, so that the researchers are becoming more interesting users of ensemble classifiers.

3. RELATED WORK

The methods of intrusion detection generally divided into three types, namely single, hybrid, and ensemble. Hybrid classifiers have the largest contribution last few years.

Koc et al. [15] proposed Hidden Naïve Bayes (HNB) with the improved naive bayes in data mining, Naïve bayes structurally extended with discretization and feature selection techniques to increase the detection accuracy. They used KDD'99 dataset for multi-class classification. HNB model significantly improved the detection accuracy of denial-of-service attacks, compared with existed models.

Wang et al. [16] proposed a new technique named FC-ANN used to improve ANN performance. Fuzzy Clustering used to generate heterogeneous training subsets, subsequently, ANN applied on each training subset to formulate the model. Meta learner and Fuzzy aggregation module is aggregate the ANN's results and reduced the detection error. This FC-ANN technique effectively detects the low-frequent attacks such as Remote to Local (R2L) and User to Remote (U2R).

Zhang et al. [17] proposed Fuzzy SVM (FSVM) to improve the payload –based anomaly detector (PAYL) which is existed. This PAYL-FSVM used reconstruction error based fuzzy membership function to reduce the data noise and solved sharp boundary

problems. PAYL-FSVM used DARPA 1999 dataset and has more powerful detection accuracy rate and low false positive rate than Payload-Based detection.

Ming-Yang Su [18] proposed a GA with KNN for feature selection and weighting. GA extracted all possible features of DoS/DDoS attacks and analyzed the relationship between performances and number of features. Detection accuracy 97.42% got when considered top 19 features for known attacks, and 78% got when considered top 28 features for unknown attacks with real-time system data.

Aydin et al. [19] designed Chaotic-based Hybrid Negative Selection Algorithm (CBHNSA) consist of generation and detection steps, negative selection and clonal selection algorithms used in the first step to obtaining optimum detectors, the detectors utilized in training stage to generate classification. CBHNSA applied both anomaly detection and classification problems. KNN used for uncovered test samples during testing to improve the detection rate.

Zhiguo et al. [20] proposed a hybrid approach of Rough Set and SVM, RS used to data reduction, SVM classified intrusions. Out of 680 data samples, 460 were applied as training samples, 220 were applied as testing samples, and results compared with traditional SVM. This technique provides better performance than the Support Vector Machine.

Giacinto et al. [21] proposed Modular Multiple Classifier System (MCS). One-class classifiers applied for designing each module, the combination of one-class classifiers distinct features considered. This MCS technique effectively combine the outputs of different classifiers with well-known fixed combination of techniques as a ensemble classifier. This model provide high detection rate and low false alarm rate with KDD CUP 99 dataset.

Tang et al. [22] designed integrated Information Gain with Triangle Area based SVM (TASVM), i.e. a combination of K-means and SVM. IG selected discriminated features, clustered the data based on 5 centroids, then calculated the triangle area using Euclidian distance to reduce the dimensionality of the feature vector, finally SVM model classified based on new feature represented triangle area. This TASVM used 10% of KDD CUP 1999 dataset, achieved 99.88% detection rate and 2.99% of false alarm rate.

Ashok et al. [23] implemented K-means Cluster Triangle Area Based SVM (CTSVM), Information Measure (IM) selected features, K-means algorithm clustered KDD data into 5 clusters, from clustered data each cluster constructed into 8 triangles, the triangle area, measured by Euclidian distance. Finally reduced the features 41 dimensions into 8 dimensions then identified the best features to identify the attack.

Amiri et al. [24] proposed Modified Mutual Information Feature Selection Method (MMIFS) for measuring the goodness of feature with the help of Linear Correlation-based Feature Selection (LCFS) and Forward Feature Selection Algorithm (FFSA). IDS model implemented with Least Squares SVM for classification. This experiment used KDD CUP 1999 dataset to detecting intrusions with high accuracy, especially for R2L and U2R attacks.

Wang et al. [25] proposed an Intelligent IDS (IIDS) included three individual IDS, used to detect WSN data threats. Intelligent Hybrid IDS (IHIDS) consist rule-based method used to identify Sinkhole attack, and has the learning capability, Hybrid IDS (HIDS) is similar to IHIDS, which identified Cluster Header (CH) attacks, but did not detect new attacks, with the help of back propagation algorithm HIDS saves resource utilization, misuse detection of HIDS retrain by IHIDS, misuse detection used to identify Sensor Node (SN) attack.

Kim et al. [26] presented a density-based outlier detection algorithm Local Outlier Factor (LOF), and proposed kd-tree indexing and approximated K-NN search algorithm (ANN) for reducing the computational time of LOF. This method significantly reduced computation time with acceptable approximation errors.

Sheikhan et al. [27] designed Feature Vitality Based Reduction Method (FVBRM) with Correlation-based Feature Selection (CFS), Information Gain (IG), and Gain Ratio (GR). Correlation-based feature selection identifies important input features. This technique used NSL-KDD dataset, Naïve Bayes approach classified the attacks with improved accuracy. FVBRM achieved more improved classification accuracy, but it takes more time.

Lee et al. [28] designed framework Self-adaptive dynamic clustering method, which is Self Organized Map (SOM) integrated with K-means clustering algorithm for real-time processing. SOM compared

given input vector with a weighted vector of each unit, and then closest unit declared as winning. K-means approach partitioned data into clusters. This method evaluated using KDD CUP 99 dataset and further used honeypot data collected from Kyoto University. The dynamic approach achieved high detection rate and low false alarm rate compared with the static model.

Sheng et al. [29] proposed combined algorithms of Partial Least Square (PLS) with Core Vector Machine (CVM), PLS extracted features and CVM process large-scale data with superior speed. PLS-CVM evaluated through KDD CUP 99 dataset. This technique PLS-CVM approach preserves the advantages of CVM modeling such as simplicity and speediness, and significantly improved the detection rate compared with CVM method.

Lin et al. [30] proposed an intelligent hybrid algorithm with Support Vector Machine, Decision Tree, and Simulated Annealing (SA). SA and SVM combined to optimize parameters and feature selection, and combined SA and DT used to build decision rules to detect new attacks. This method evaluated using KDD CUP 99 dataset, and outperform than existing approaches.

Gaikwad et al. [31] designed Fuzzy Clustering ANN (FC-ANN), fuzzy clustering generated heterogeneous training subsets, subsequently different ANN models formulate the models with training subsets. Fuzzy aggregation module used to aggregate different ANNs results and detection errors reduced. The experiment evaluated using KDD CUP 1999 dataset and demonstrated the effective results especially for low-frequent attacks R2L and U2R.

Shin et al. [32] proposed Adaptive Probabilistic Approach for Networks (APAN), which is K-means clustering defined network states and introduced outlier factor concept. A Markov model included state transition probability matrix and built initial probability distribution. In real-time, the degree of incoming data abnormality measured stochastically using the model. The performance of this method evaluated using DARPA 2000 dataset. This method is insensitive to training dataset variations and number of Markov model states.

Aneetha and Bose [33] proposed network anomaly IDS with combined Modified SOM and K-means, MSOM dynamically created new nodes with help of distance threshold, connection strength, and neighborhood function. K-means clustering algorithm grouped similar nodes of MSOM into k number of clusters using with a similarity measure. This MSOM improved 2% detection rate higher compared with existed SOM, it is further increased by 1.5% when applied K-means. It effectively detected the DDoS attacks with 98.5% detection rate.

Ibrahim et al. [34] multi-layer IDS designed to attain high detection rate and classification accuracy with Decision Tree, Multi-Layer Perceptron, and Naïve bayes. Gain Ratio (GR) extracted features, C5 DT achieved high classification rate and less false alarm rate achieved by MLP and Naïve Bayes. The experiment evaluated using NSL-KDD dataset. With the help of Gain Ratio increased the detection accuracy of less frequent attacks R2l and U2R. MLP has high rate of classification for DoS and Probe.

Kim et al. [35] designed C4.5 Decision Tree algorithm with Multiple One-Class SVM. C4.5 decision Tree, decompose the normal training data into subsets, then multiple OCSVM used to create anomaly detection model on each decomposed misuse region. This technique evaluated using KDD CUP 99 dataset, and demonstrates the better detection rate compared to conventional methods for both known and unknown attacks.

Chandrashekhar and Raghuveer [36] proposed a hybrid IDS technique with Fuzzy C-means Clustering algorithm. FCM clustered the input data, each of the data point associated with cluster trained by FNN (neuro-fuzzy) classifier, subsequently, SVM with Radial Basis Function (RBFSVM) applied for classification to detect an anomaly or normal. Experiment evaluated using KDD CUP 1999 dataset, obtained better results for all metrics.

Zainaddin and Hanapi [37] proposed a hybrid framework clustered the data by Fuzzy Clustering and classified the attack type by ANN. This framework tested with both KDD CUP 1999 and NSL-KDD dataset. The results of precision, f-value rate and recall are compared with the previous experiment. Both datasets cover the main types of attacks effectively.

Thaseen and Kumar [38] proposed a new technique integrating PCA and SVM adopted RBF kernel,

parameters obtained using a grid search with automatic parameter selection. PCA select the optimal subset of attributes. The method tested with KDD CUP 1999 dataset, and attained better accuracy for low-frequency attacks U2R and R2L. The classification accuracy of this approach outperforms than other classification techniques using SVM.

Chitrakar and Huang [39] proposed a technique which contributed in two ways, first Candidate support vector incremental SVM classification (CSV-ISVM) that is obtained by modified existed concentric-ring method and reserved set strategy, secondly efficient and faster Support vector selection method named as Half-partition strategy to reduce classification time along with high detection rates. This method evaluated using KDD CUP 1999 dataset and Kyoto 2006+ dataset and attained better detection rate with a low false alarm rate and acceptable learning time.

Desale and Ade [40] proposed GA based feature selection method used different feature selection techniques Correlation based Feature selection (CFS), Information Gain (IG), and Correlation Attribute evaluator (CAE) for feature selection, and performance classified by Naïve bayes or J48. This method used NSL-KDD dataset and selected the minimum number of features, and improved accuracy of Naïve Bayes classifier with reduced time.

Mohd Pozi et al. [41] proposed a hybrid model to detect anomalous rare attacks by combining SVM and GP named GPSVM, GP is an optimization technique used to solve problems, SVM used as a classifier to detect attacks, meanwhile, DT used to reduce the GPSVM decision function complexity. This technique produced more balanced accuracy on NSL-KDD dataset without using any reduction method such as feature selection and feature extraction.

Jabez and Muthukmar [42] designed new technique outlier detection, anomaly dataset measured by Neighborhood Outlier Factor (NOF), SNORT used to collect datasets and extract the features, designed IDS compute distance between the extracted features and trained model, NOF used to detect outliers, improved the performance of IDS with distributed storage environment of big dataset. This method tested with KDD CUP 1999 dataset and attained better performance in terms of detection rate than existed machine learning approaches.

Ikram and Aswani Kumar [43] proposed IDS model integrating of PCA and SVM using RBF kernel. PAC reduced the dimensionality of data, SVM classification model constructed based on training data obtained from PCA. SVM parameters C and ϒ optimized for RBF kernel by existed automatic parameter selection technique to classification model. This experiment evaluated with two different datasets NSL-KDD and gurekddcup dataset and obtained results in terms of classification accuracy.

Azad and Jha [44] proposed Fuzzy min-max neural network and Particle Swarm Optimization, the learning is executed by a series of hyper box expansion, the hyper box expansion depends on box size. PSO optimize hyper box min max values and classify the attacks. This technique evaluated using KDD CUP 1999 dataset and attained greater performance compared with MLP.

Salunkhe and Mali [45] proposed ensemble classifier, combined different base classifiers, the data detection model 1 extracted original dataset and created data subsets model 2 extracting original training data feature subset and created feature subset, then combined outputs of two models obtained the final prediction by ensemble classifier whether intrusion exist or not. This technique evaluated on KDD CUP 1999 dataset, performance tested for more number of attack categories.

Mabu et al. [46] proposed traditional GNP based rule mining method and evolutionary optimization technique, generates rule database and classifies new data with ensemble Random Forest. GA optimizing the weights of training dataset classifies the test data by integrating the classification results using weighted majority vote of rule databases. This technique used NSL-KDD dataset, and obtained better detection accuracy compared with SVM and J4.8.

4. FINDINGS OF THE LITERATURE SURVEY

In this area, let us discuss the different techniques utilized by various authors in their research works. From the literature review discussed above, it is concluded that most of the researchers used classifiers and cluster methods as an Intrusion Detection System in earlier. Later feature extraction methods are used to

extract the important features, along with classifiers. In this concerned area, many researchers used hybrid classifiers for Intrusion Detection, which are a combination of feature extracted, clustering, and classification methods. Very few numbers of researchers used ensemble classifiers as Intrusion Detection techniques. Most of the proposed techniques evaluated on KDD CUP 1999 dataset, and a revised version of KDD'99 dataset is NSL-KDD dataset using recently by the researchers. After proper analysis of above-discussed literature and techniques, we concluded that.

- In recent years many of researchers have designed hybrid classifier technique with feature selection methods. These methods attained higher accuracy compared with single classifiers. On the other hand small amount of work done on ensemble classifiers.
- The feature selection technique Information Gain (IG) with classification techniques KNN, NN, NB etc. produced higher performance compared with other feature selection techniques, on the other hand, Correlation Feature selection (CFS) with DT and Principle Component Analysis (PCA) with SVM produced higher performance.
- Majority of the techniques used benchmark datasets such as KDD CUP 1999 dataset and NSL-KDD dataset. Approximation of the actual performance of the intrusion detection in real time data is difficult to evaluate.
- Recent researchers have been applied data mining and machine learning techniques. Most of these techniques were built on shallow learning architectures. These architectures are still somewhat unsatisfying for intrusion detection.
- Many researchers have classified 2- class classifications either as an attack or as normal.
- The classical machine learning algorithms are inefficient to detect zero-day attacks, and low frequency attacks such as R2L and U2R due to insufficient quantity of labeled training data and network traffic variability.

5. CONCLUSION

We have reviewed current studies of Machine Learning Intrusion Detection techniques. This study included a single, hybrid, and ensemble classifiers. Building an effective intrusion detection system using Machine learning and Deep learning methods have received much attention for network security. Data set always contain a huge number of features where most of them are redundant or irrelevant. Employing a feature reduction method is essential to reduce the computational cost and increase the classifier performance. Feature selection and feature extraction are having advantages and useful to detect attacks with classifiers, which make it hard with a single classifier method to implement. It's recommended to use feature extraction followed by feature selection as a hybrid approach to increase the accuracy of intrusion detection.

It is there by recommended for an improvement of intrusion detection in the field of cyber security, and thus Deep Learning algorithms are more adaptable systems on systems with faster processing capability with Graphical Processing Unit (GPU).

REFERENCES

[1] Stallings, W. (2006). Cryptography and network security principles and practices. USA:Prentice Hall.

[2] Anderson, J. (1995). An introduction to neural networks. Cambridge: MIT Press.

[3] Karen Scarfone ,Peter Mell "Guide to Intrusion Detection and Prevention Systems (IDPS)", National Institute of Standards and Technology Special Publication 800-94, 2007

[4] S. Manocha , M.A. Girolami, "An empirical analysis of the probabilistic K-nearest neighbour classifier", Science Direct, Pattern Recognition Letters, 28 (2007), 1818–1824.

[5] C.M.Bishop. (1995). Neural networks for pattern recognition. England: Oxford University.

[6] Mitchell, T. (1997). Machine learning. New york: MacHraw Hill.

[7] Chih-Fong Tsai, Y.-F. H.-Y.-Y. (2009). Intrusion detection by machine learning: A review. expert systems with applications, ELSEVIER

[8] J. R. Quinlan, "Introduction of Decision Trees", Machine Learning, Centre for Advanced Computing Sciences, New South Wales Institute of Technology, Sydney ,2007, vol. 1

[9] S. Haykin, Neural networks: A comprehensive foundation (2nd ed.), Prentice Hall, New Jersey, U.S.A, 1999.

[10] Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. Biological Cybernetics, 43, 59–69.

[11] Koza, J. R. (1992). Genetic programming: On the programming of computers by means of natural selection. Massachusetts: MIT.

[12] Pearl, J. (1988). Probabilistic reasoning in intelligent systems. Morgan Kaufmann.

[13] H. Zimmermann, Fuzzy set theory and its applications. Kluwer Academic Publishers

[14] J. -S. Jang, C. -T. Sun, and E. Mizutani, Neuro-fuzzy and soft computing: A computational approach to learning and machine intelligence. Prentice Hall, New Jersey, USA, 1996.

[15] Levent Koc, Thomas A. Mazzuchi, Shahram Sarkani," A network intrusion detection system based on a Hidden Naïve Bayes multiclass classifier ", Expert Systems with Applications, Vol.39,pp. 13492–13500,2012.

[16] Gang Wang, Jinxing Hao, Jian Ma, Lihua Huang," A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering ", Expert Systems with Applications, Vol.37, pp. 6225–6232, 2010.

[17] Guiling Zhang, Yongzhen Ke, Liankun Sun, Weixin Liu," An Improvement of Payload-based Intrusion Detection Using Fuzzy Support Vector Machine", 2nd International Workshop on Intelligent Systems and Applications, IEEE,2010.

[18] Ming-Yang Su,"Real-time anomaly detection systems for Denial-of-Service attacks by weighted k-nearest-neighbor classifiersr", Expert Systems with Applications, Vol.38, pp.3492–3498,2010.

[19] Ilhan Aydin, Mehmet Karakose, Erhan Akin," Chaotic-based hybrid negative selection algorithm and its applications in fault and anomaly detection", Expert Systems with Applications, Vol.37, pp.5285–5294, 2010.

[20] Liu Zhiguo, Kang Jincui, Li Yuan, "A Hybrid Method of Rough Set and Support Vector Machine in Network Intrusion Detection", 2nd International Conference on Signal Processing Systems, pp.561-563.2010.

[21] Giorgio Giacinto, Roberto Perdisci, Mauro Del Rio, Fabio Roli, "Intrusion detection in computer networks by a modular ensemble of one-class classifiers", Information Fusion Vol. 9 pp. 69–82, 2008.

[22] Pingj ie Tang, Rang-an Jiang, Mingwei Zhao, "Feature selection and design Of intrusion detection system based on k-means and triangle area support vector machine", Second International Conference on Future Networks,pp.144-148,2010.

[23] Ashok R, A Jaya Lakshmi, G Devi Vasudha Rani, Madarapu Naresh Kumar, "Optimized Feature Selection with k-Means Clustered Triangle SVM for Intrusion Detection", Third International Conference on Advanced Computing, 2011.

[24] Fatemeh Amiri, MohammadMahdiRezaeiYousefi , CaroLucas , AzadehShakery, NasserYazdani,"Mutual information-based feature selection for intrusion detection systems", Journal of Network and Computer Applications, Vol.34, pp.1184–1199,2011.

[25] Shun-Sheng Wang, Kuo-Qin Yan, Shu-Ching Wang, Chia-Wei Liu," An Integrated Intrusion Detection System for Cluster-based Wireless Sensor Networks ", Expert Systems with Applications,Vol.38, pp. 15234–15243, 2011.

[26] Seung Kim, Nam Wook Cho, Bokyoung Kang, Suk-Ho Kang," Fast outlier detection for very large log data ", Expert Systems with Applications, Vol.38, pp. 9587–9596, 2011.

[27] Dr. Saurabh Mukherjee, Neelam Sharma," Intrusion Detection using Naive Bayes Classifier with Feature Reduction ", SciVerse Science Direct,Vol. 4,pp.119-128, 2011.

[28] Seungmin Lee, Gisung Kim, Sehun Kim," Self-adaptive and dynamic clustering for online anomaly detection", Expert Systems with Applications,Vol.38, pp.14891–14898,2011.

[29] Gan Xu-sheng, Duanmu Jing-shun , Wang Jia-fu , Cong Wei," Anomaly intrusion detection based on PLS feature extraction and core vector machine ", Knowledge-Based Systems, Vol.44, pp.1–6,2012.

[30] Shih-Wei Lin, Kuo-Ching Ying, Chou-Yuan Lee, Zne-Jung Lee, "An intelligent algorithm with feature selection and decision rules applied to

anomaly intrusion detection", Applied Soft Computing, Vol.12,pp. 3285-3290, 2012.

[31] Prof. D.P. Gaikwad, Sonali Jagtap, Kunal Thakare, Vaishali Budhawant, "Anomaly Based Intrusion Detection System Using Artificial Neural Network and Fuzzy Clustering", International Journal of Engineering Research & Technology (IJERT),Vol.1, pp.1-7, 2012.

[32] Seongjun Shin, Seungmin Lee, Hyunwoo Kim, Sehun Kim, "Advanced probabilistic approach for network intrusion forecasting and detection", Expert Systems with Applications, Vol.40, pp. 315–322, 2012.

[33] A.S. Aneetha and Dr. S. Bose," THE COMBINED APPROACH FOR ANOMALY DETECTION USING NEURAL NETWORKS AND CLUSTERING TECHNIQUES", Computer Science & Engineering: An International Journal (CSEIJ), Vol.2, No.4,pp.37-46,2012.

[34] Heba Ezzat Ibrahim, Sherif M. Badr, Mohamed A. Shaheen, "Adaptive Layered Approach using Machine Learning Techniques with Gain Ratio for Intrusion Detection Systems ", International Journal of Computer Applications,Vol.56, pp.10–16, 2012.

[35] Gisung Kim, Seungmin Lee, Sehun Kim," A novel hybrid intrusion detection method integrating anomaly detection with misuse detection ", Expert Systems with Applications, Vol.41,pp. 1690–1700, 2013

[36] A. M. Chandrashekhar and K. Raghuveer," Fortification of Hybrid Intrusion Detection System using Variants of Neural Networks and Support Vector Machines ", International Journal of Network Security & Its Applications (IJNSA), Vol.5, No.1, pp.71-90, 2013.

[37] Dahlia Asyiqin Ahmad Zainaddin and Zurina Mohd Hanapi," Hybrid of Fuzzy Clustering Neural Network Over NSL Dataset for Intrusion Detection System ", Journal of Computer Science, Vol.9, pp.391-403, 2013.

[38] I. Sumaiya Thaseen, Ch.Aswani Kumar, "Intrusion Detection Model Using fusion of PCA and optimized SVM", International Conference on Contemporary Computing and Informatics (IC3I), 2014.

[39] Roshan Chitrakar, Chuanhe Huang," Selection of Candidate Support Vectors in incremental SVM for network intrusion detection", ScienceDirect, pp.231-241,2014.

[40] Mr. Ketan Sanjay Desale, Ms. Roshani Ade," Genetic Algorithm based Feature Selection Approach for Effective Intrusion Detection System", International Conference on Computer Communication and Informatics,pp.1-6,2015.

[41] Muhammad Syafiq Mohd Pozi, Md Nasir Sulaiman, Norwati Mustapha, Thinagaran Perumal, "Improving Anomalous Rare Attack Detection Rate for Intrusion Detection System Using Support Vector Machine and Genetic Programming", Neural Process Lett Vol. 44, pp. 279–290, 2015.

[42] JABEZ J, Dr.B.MUTHUKUMAR,"Intrusion Detection System (IDS): Anomaly Detection using Outlier Detection Approach", International Conference on Intelligent Computing, Communication & Convergence, 2015.

[43] Sumaiya Thaseen Ikram and Aswani Kumar Cherukuri," Improving Accuracy of

[44] Intrusion Detection Model Using PCA and Optimized SVM ", Journal of Computing and Information Technology, Vol. 24, pp.133–148,2016.

[45] Chandrashekhar Azad, Vijay Kumar Jha," Fuzzy min–max neural network and particle swarm optimization based intrusion detection system", Microsyst Technol, 2016.

[46] Uma R. Salunkhe, SureshN. Mali, "Security Enrichment in Intrusion Detection System Using Classifier Ensemble", Journal of Electrical and Computer Engineering, 2017.

[47] Shingo Mabu, Shun Gotoh, Masanao Obayashi, Takashi Kuremoto, "A random-forests-based classifier using class association rules and its application to an intrusion detection system", Artificial Life and Robotics , Vol.21, pp.371-377, 2016.