Advances in Computer Science and Information Technology (ACSIT)
Print ISSN: 2393-9907; Online ISSN: 2393-9915; Volume 2, Number 8; April-June, 2015 pp. 77-81

http://www.krishisanskriti.org/acsit.html

A Summarization of Various Intrusion Detection Techniques

Randeep Brar[1] and Neeraj Sharma[2]

[1]Deptt. of Computer Science and Engineering Chandigarh Engineering College Landran, Mohali, India
[2]Head of Department(CSE) Chandigarh Engineering College Landran, Mohali, India
E-mail: [1]randeepbrr8@gmail.com, [2]hodcse@cecmohali.org

Abstract—*Intrusion detection is one of the most discussed topics in the network security. Also, immense increase in authorized access and network attacks has enforced the new and tough challenge for intrusion detection methods. The elevated amount of the data on the network also feels the necessity of an efficient technique to carry out intrusion detection in rapid time. Numerous IDS techniques have been considered so far, which differ from each other on the account of various factors. Current IDSs face challenges not only on low detection rates but also on huge consumption of training data and time. This paper contains the summarization study of all the techniques and identification of pros and cons of all the techniques, which are formerly surveyed. Among all the techniques reviewed, some techniques give the optimistic result while some techniques are still in the need of improvement. Hitherto, the techniques to carry out intrusion detection are boundless, but they all lag behind in one feature or the other. The proper blend of the useful techniques can result in an ideal or near to ideal intrusion detection system*

1. INTRODUCTION

The development of World Wide Web and increase in amount of data exchange has raised an issue of security over the network and on the system. Every system should provide confidentiality against the various attacks or intrusions. Intrusion can be said as any premeditated action that results in unauthorized access or instability in the system such as DOS attacks or information disclosure. Intruders are the users who carry out the intrusion i.e. any action resulting in attack. Intrusions may result as a heavy loss for the organizations, whose system bears the attack. Sometimes the organizations have to bear the loss of million dollars. So to prevent all these intrusions detection system is developed.

Intrusions detection system can be defined as a system used to recognize attack against the system (host) or the network. Intrusions detection can be classified on the basis of the analysis technique used and on the basis of the positioning of the intrusions detection system. When the IDS is classified on the basis of analysis technique, then it can divided as misuse detection and anomaly detection. In misuse detection, IDS uncovers the known attack i.e. the attack whose signatures (patterns) are already stored in the data base i.e. the IDS cannot expose the 'unseen'. In anomaly detection, IDS detect the unlabelled attacks i.e. it detects the new types of attacks. On the basis of the positioning of the IDS it is split as host-based or network-based systems. The host-based IDS detects the attacks on the same system where it is positioned .The network-based IDS expose the malicious activities or traffic on the network. Various techniques can be used to improve the intrusion detection method. Some techniques aim to reduce training data and time by carrying out feature selection whether some techniques alter the classification method for easy generalization. Numerous clustering techniques are used to improve the intrusions detection system. Clustering aims to combine the data and the clusters are formed on the basis of degree of similarity and dissimilarity of the data objects. Clustering uses an unsupervised learning mechanism. Unsupervised learning mechanism finds the different pattern among the unlabelled data. There are various factors for the clustering process. The clustering can be done using the distance as the similarity measure, whereas some use density based clustering method or other may use grid-based clustering algorithm. The clustering technique may have impact on the detection rate. Some of the recent researchers say that including an intelligent clustering algorithm such as PSO may favor the IDS result.

The importance of IDS can known from the CISCO [14] website. According to the website [14] Cisco internal Information Security (Infosec) group is responsible for security within the Cisco internal network. Infosec deployed more than 35 Cisco IDS 4230 sensors worldwide in 2001, and receives alerts about the malicious activities in the internal network of organization. IDS also traces the pattern and port used for attack. This helps Infosec to act rapidly to an intrusion and to secure the systems against an attack with same pattern in the future. For example, in 2001 there were two disastrous attacks on servers worldwide: Code Red 1 and 2 (July 2001) and NIMDA (September 2001) [14]. Both times the Cisco IDS 4230 sensors noticed an unusually sharp rise in malicious traffic [14] within the first few minutes, and within an hour Infosec and IT engineers took immediate step to curb the menace. If the IDS would not be used by Infosec, it is possible that server would have been affected badly, CISCO

would have more financial loss and it would be a source of far more infection over the internet. Another SQL slammer alert was sent to entire industry and IDS was used to secure the system and Cisco remained intact by the attack.

In this paper, we have tried to analyze the various technique proposed by researchers. The analyzed result concludes some most successful techniques, which have the future scope.

Paper organization: section 2 includes the related work and section 3 give the concluding remark of the revision

2. LITERATURE REVIEW

The intrusion detection includes various steps and the minor change in any of the step may vary the result. The researchers worldwide has tried alteration in the techniques and compared the result. Among all those researchers, authors in [9] proposed a PSO based clustering method IDCPSO which combines an unsupervised clustering algorithm with PSO. The IDCPSO clustering has two stages, the stage of clustering and stage of PSO optimization. The clustering is based on similarity among the instances. In PSO optimization, the initial clusters set the initial particle and the PSO procedure improves the swarm repeatedly, due to collective intelligence of particles and this procedure goes on till the fitness value does not stop changing. After this procedure, the clusters are labeled as normal and 'anomalous'. Anomaly detection follows the procedure of clustering. The KDD cup' 99 data set is used for the experiment. The evaluation of the data set is carried on the basis of the DR(Detection Rate), FPR(False Positive Rate), FNR (False Negative Rate) . The performance of the proposed algorithm is compared with Genetic Algorithm (GA). The comparison result shows that using IDCPSO results in higher detection rate and low FPR and low FNR. The result also highlights that speed of convergence process is higher in proposed algorithm than Genetic algorithm. The future of this technique lies in the fact that detection rate is improved.

Additional method which preferred PSO based clustering is introduced by Aljarah et al. [1] using the MapReduce methodology. The idea introduced by them is to use particle swarm optimization for data mining as PSO avoids the sensitivity problem of initial cluster centroids [1]. This model processes the data in bulk amount. The similarity measure is the distance as distance is considered as important factor to classify the object. The parallel intrusion detection system (IDS-MRCPSO) uses the MapReduce methodology [1]. This methodology favors parallelization of data mining tasks. Performance is measured using the parallel speedup [1]. The quality of intrusion is checked in TPR, FPR etc. The ROC curve evidences the success of the technique. The results show that methodology used by them justifies the detection in large amount of data and the result lies close to optimal speedup. Also, instead of considering some parts of data set, the whole training set is used to build the detection model. Hence, results into an improved detection model i.e. the no. of false alarms raised are very low. The proposed model (IDS-MRCPSO) can be further improved by adding the factor to distinguish between different types of intrusions.

One more clustering technique, 2-tier clustering technique [13] is used by the paper entitled 'An adaptive approach of clustering application in the intrusion detection systems. The proposed model is based on the classification trees and distance based comparison of system call sequences. This methodology is designed in aim to trace attack during current activities in system. The proposed method divides the data into two non- intersecting clusters. One cluster contains normal pattern and other contains anomalous pattern. The object's clustering is done on the basis of the distance. The evaluation of the performance is done on the four point's basis: true positive (TP), true negative (TN), false positive (FP), false negative (FN). For quality evaluation, the rand index is used. The rand index calculates the degree of similarity between the clusters.

Leung et al. [8] used the (fpMAFIA) density based and grid based clustering algorithm for unsupervised anomaly detection in network intrusion detection. The unsupervised anomaly detection techniques make the system capable of detecting unseen attacks. They made two assumptions to use unsupervised anomaly detection algorithm. They assumed that majority of network connection are normal and only X% are malicious. Their second assumption is that statically, attack traffic is different from normal traffic. They used density based method [8] and grid-based method [8] for clustering. The density based methods grow the clusters till the objects do not increase the threshold. In grid based method, the clustering operations are performed on objects, which are arranged into finite no. of cells. The fpMAFIA can produce clusters of any shape and may cover much of the data sets. The algorithm used here is adapted from other optimized algorithm with some modifications. The fpMAFIA runs with large data set of 1 million records on a single system. After obtaining the clusters, the sub- clusters or anomalies are also removed. The data set used for experiment is a filtered 1999 KDD data set .The filtration is carried out to satisfy the assumptions. After training phase, the set of the clusters are obtained and the objects outside the clusters are classified as anomalies. The overall evaluation shows that fpMAFIA algorithm has resulted in high detection rate but high false positive rate.

Another approach proposed by Sharma et al. [2] for intrusion detection is based on k-Means clustering via Naïve Bayes Classification [2]. They have considered network intrusion detection system to monitor the network activities. The comparison is carried out on two techniques for intrusion detection: i) on k-Means clustering via Naïve Bayes Classification ii) Naïve Bayes Classification. The intrusion detection technique proposed in [2] uses KDD data set to carry out experiments. The basic step of preprocessing and normalization is carried out. Then the k-Means clustering focuses on grouping objects based on feature values. Based

on these steps, detection rate and false positive rate is given. The comparison between the Naïve Bayes Classification and k-Means clustering approach shows that k-Means clustering has 99% detection rate with 4% error rate whereas Naïve Bayes Classification has 95% detection rate. The limitation of the proposed model is that it generates more false positives.

All the techniques in [9], [1], [13], [8], [2] carried out alteration in the clustering. In [9], PSO algorithm is used in clustering whereas in [1] researchers use PSO combined with MapReduce methodology. Researchers in [13] used 2-tier distance based clustering while in [8] density and grid based method was preferred. K-Means clustering is combined with Naïve Bayes Classification in [2] and result is compared with Naïve Bayes Classification. All techniques have their own results.

The neural network based intrusion detection is proposed for years. The MLP neural networks [5] are used for detection. For detection, the traffic is grouped together and clustering is carried out. The data used is both offline and online type. The system reads the data and then passes it on for three steps: preprocessing, clustering and normalization. The paper laid emphasis on the fact that before carrying out the monitoring of network traffic, the neural network structure should be determined. The paper [5] has given the reason for the energy step carried out during the research procedure. The paper concludes that using a neural network for intrusion detection gives the promising result. The clustering method, SOMs[5], for MLP (Multilayer Perceptron) neural network is a promising way of creating perfect and grouped input for detection, for dynamic no. of inputs. The future scope for this research is use of both supervised and unsupervised learning analysis of network traffic.

One more point of view of use of MLP neural network is elaborated by Moradi and Zulkerine [4]. A neural network based system for classification and detection of attack [4] is proposed. An offline IDS is implemented using MLP (Multilayer Perceptron) [4] artificial neural network. This research excels by the fact that it not only aims to spot the intrusion but also classifies the attacks. The classification enables the proper action against the specific type of attacks. Thais leads to the development of practical intrusion detection system. DARPA dataset is used for the evaluation. Due to validation method used, the training time is decreased and generalization capacity is increased. For the classification of connection records, two layer neural networks or three layer neural network is used. Although the three layer network performs better, but for the research work two layer network is preferred as it is less complicated, saves money and is computationally more efficient. The experiment results show that on training set, 93% classification result is correct and in testing set, 87% classification result is correct.

The research in [4] could be further improved by more attack scenario in the dataset. Also, the complexity in the network could be improved by dividing connection record to normal and general categories of record and then further classifying into attack types.

A new learning methodology is proposed by Chang et al. [7]. They develop an intrusion detection system by back propagation neural networks with sample query and attribute query [7]. The proposed learning methodology in paper aims to improve the performance as this method first selects good attributes by applying information theory. Then query based methodology includes the subsets of samples in learning. This method leads to accurate prediction of portable attack behavior in IDS. The proposed method may explore the unseen attacks. The experiment is carried out on the KDD data set. The test data is not from the same probability distribution as training data [7]. The comparison of BPN neural network and BPN neural network with sample query and attribute query is done. The result of the comparison is presented using confusion matrix [7] which shows training time of proposed system is very less than training time of BPN.

Another classification method used is LSSVM (Least square Support Vector Machine) in intrusion detection system in [12]. The computation method used is chaos particle swarm optimization. LSSVM is based on structure risk minimization [12]. CPSO [12] is responsible for optimal settings of LSSVM parameters. The paper proposes the combination of CPSO-LSSVM. The overall results show that CPSO-LLVM has better detection rate as compared to BPNN (Back Propagation Neural Network).

The intrusion detection can be improved by the effective feature selection. The researchers in [10] show that selecting the best features among all the features result in improved detection rate. According to the paper [10], the redundant features affect the use of the system resources and time. Feature selection model using data mining is proposed to overcome this problem. The irrelevant features slow down the testing time and training time and consumes much resources. The proposed model contains the four stages- data processing, best classifier selection, feature reduction and best feature selection. Feature reduction removes the redundant features. The best feature selection has two phases on the basis of feature set chosen to be start set of the phase. To check the effectiveness of the proposed model, various performance measures are used such as specificity, sensitivity, TPR, FPR and classification accuracy etc. [10]. The proposed model with best 11 features is compared with 41 feature set. The proposed model is more time efficient and has more detection rate than one with complete feature set. However, few probe attack instances are misclassified as DOS attacks by proposed model.

One more perspective of intrusion detection modeling is given by More et al. in [11] is knowledge based approach. The work proposed an intrusion detection model which is not only a signature based system, but it is a situation- aware intrusion detection [11] model that uses heterogeneous data sources, so that rich knowledge base is build which can recognize cyber

crimes. The proposed model overcomes the limitation of the signature based IDS, as the latter cannot distinguish the attack whose signature is not stored in their database i.e. have 'unseen' attacks. The proposed model [11] has information module, data from different sensor streams, and text data from web etc as key components. The data stream includes data from different sources and the items of interest are added to the knowledge base. The working of the system is checked by simulating the attack in controlled environment on local network, but the IDS scanner does not trace the attack, since it is not stored in the database. Thus, the strong knowledge base enables the IDS to spot the variety of attacks.

Omar Al-jarrah & Ahmed Arafat [3] chose network intrusion detection system to carry out research work & used attack behavior classification for it. The paper showed the port attacks. Port attacks discover various services available on the network. For data processing, neural networks are used. The intelligent system is used to carry out detection process for stored signature in efficient manner. This technique is mainly used to increase detection rate & decrease false positive rates. Data is captured in real time using packet capture engine. The processing stage, feature selection is carried out by using TDNN (Time Delay Neural Network)[3] .The pattern recognition components trace attack by using the neural networks which carries out PCA (Principal Component Analysis)[3]. The classifier then classifies the attack and generates alert. The DARPA dataset is used to evaluate the proposed system. While testing, the system is compared to rule based IDS SNORT [3]. The test result shows that the proposed system detects more attacks than SNORT.

Despite the plethora techniques of intrusion detection, the network security possesses a threat. Some of the intrusion detection techniques are reviewed by Vinchurkar and Reshamwala [6]. In their paper they have given basic structure of the intrusion detection system and also describe the ideal intrusion detection system. They have categorized intrusion detection system on the basis of data source and model of intrusion. Various challenges in intrusion detection are well presented. Neural network approach and machine learning approach for intrusion detection are evaluated. They also explain dimension reduction using PCA [6]. They reach the conclusion that the current IDS is effective to update audit data fast and also, there is need to design the IDS which can compete with the current challenges of holding bulky database and improving performance measures.

3. CONCLUSION

Numerous intrusion detection techniques are reviewed in the paper. The author of every paper has carried out alteration in one or more steps in the intrusion detection system to raise the detection rate and to lower the false results. Each intrusion detection technique, reviewed in the paper has pros and cons. To declare the best among those is quite impossible, as a technique may lag in one feature and excel in other.

Many of the papers have preferred alteration in clustering as clustering is the backbone of the whole procedure while some papers preferred to filter feature selection. Among all the papers reviewed, a few papers favored use of neural networks for data processing and classification etc.

Besides detecting signature based attack, an intrusion detection system should also detect the unknown attacks. The intrusion detection system should even trace the minor malicious activities over the network or in the system

In addition to all the forgoing, it is concluded that combining the best techniques and eliminating the weakness can result as ideal IDS which justifies the challenges in internet and data security.

REFERENCES

[1] Ibrahim Aljarah and Simone A. Ludwig "MapReduce Intrusion Detection System based on Particle Swarm Optimization Clustering Algorithm", *IEEE Congress on evolutionary Computation, June 20-23, 2013*

[2] Sharma S. K., Pandey P., Tiwari S. K., Sisodia M. S., "An Improved Network Intrusion Detection Technique based on k-Means Clustering via Naïve Bayes Classification *",Advances in Engineering, Science and Management (ICAESM), 2012 International Conference on [proceedings] : date, 30-31 March 2012*. Piscataway, NJ: IEEE, 2012.

[3] Omar Al-Jarrah & Ahmed Arafat, " Network Intrusion Detection System Using Attack Behaviour Classification",5th *International Conference on Information and Communication Systems (ICICS),2014*

[4] Moradi , M., Zulkernine, M.,"A Neural Network Based System for Intrusion Detection and Classification of Attacks", *Natural Sciences and Engineering Research Council of Canada (NSERC).*

[5] Alan Bivens, Mark Embrechts, Chandrika Palagiri, Rasheda Smith, and Boleslaw Szymanski, "Network-Based Intrusion Detection Using Neural Networks", *Artificial Neural Networks In Engineering, St. Louis, Missouri, November 2002*.

[6] Alpa Reshamawala and Deepika P. Vinchurkar, "A Review of Intrusion Detection System using Neural Network and Machine learning Technique" *International journal of Engineering Science and Innovative Technology, Vol.1, Issue 2, November 2012*

[7] Chang, Ray-I, Liang-Bin Lai, Wen-De Su, Jen-Chieh Wang, and Jen-Shaing Kouh. "Intrusion Detection by Back propagation Neural Networks with Sample-Query and Attribute-Query", *International Journal of Computational Intelligence Research, Vol. 3, No.1, 2007.*

[8] Kingsly Leung and Christopher Leckie, "Unsupervised Anomaly Detection in Network Intrusion Detection Using Clusters", *Australasian Computer Science Conference, Newcastle, NSW, Australia, 2005.*

[9] Zheng, Hongying, Meiju Hou, and Yu Wang. "An Efficient Hybrid Clustering-PSO Algorithm for Anomaly Intrusion Detection." *Journal of Software Vol. 6, December 2012.*

[10] Ayman I. Madbouly, Amr M. Gody, Tamer M. Barakat, "Relevant Feature Selection Model Using Data Mining for Intrusion Detection System", *International Journal of*

Engineering Trends and Technology (IJETT) Vol. 9, No. 10, March 2014.

[11] Sumit More, Maru Mathews, Anupam Joshi Tim Finin, "A Knowledge-Based Approach To Intrusion Detection Modeling", *IEEE CS Security and Privacy Workshops, 2012.*

[12] Mohua Zhang and Ge Li, "Network intrusion detection based on least square support vector machine and chaos particle swarm optimization algorithm", *Journal of Convergence Information Technology, Vol. 7, No. 4, March 2012.*

[13] Evgeniya Petrova Nikolova & Veselina Gospodinova Jecheva "An Adaptive Approach of Clustering Application in the Intrusion Detection Systems", *Open Journal of Information Security and Applications, Vol. 1, No. 3, December 2014.*

[14] Security Case Study: How Cisco IT Upgraded Intrusion Detection to Improve Scalability and Performance<http://www.cisco.com/web/about/ciscoitatwork/security/intrusion_detection_upgrade_web.html>