International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 26 Issue 6 – OCTOBER 2019.

Performance comparison of support vector machine, random forest, and extreme learning machine for intrusion detection

S.Kavin & S.Krishna Mohan & V.I.Karthick [#1] and K.Muthamil Sudar [*2]

[#] *Department of Computer Science Engg., Kalasalingam Academy of Research & Education, Krishnankovil, India*

[*] *Department of Computer Science Engg., Kalasalingam Academy of Research & Education, Krishnankovil, India*

***Abstract*— Intrusion detection is a basic part of security tools, such as adaptive security appliances, intrusion detection systems, intrusion prevention systems and firewalls. A new technology MANET have an important strength to be applied in critical situations like battlefields and commercial applications such as building, traffic surveillance, MANET is infrastructure less, without some centralized controller exist and also each node contain routing capability, Every device in a MANET is independently free to move in any direction, and will therefore change its connections to other devices frequently. So one of the main challenges wireless mobile ad-hoc networks face today is security, for the reason no central controller exists. MANETs are an one kind of wireless ad hoc networks that usually has a routable networking environment on top of a link layer ad hoc network. Ad hoc also contains wireless sensor network so the problems are facing by sensor network is also faced by MANET. While creating the sensor nodes in unattended environment increases the chances of various attacks. There are several security attacks in MANET and DDoS(Distributed denial of service) is one of them. Our main goal is seeing the effect of DDoS in routing load, packet drop rate, end to end delay, i.e. maximizing due to attack on network. And based on these parameters and many more also we build secure IDS to detect this kind of attack and block it. In this journal, we discussed some attacks on MANET and DDOS also and provide the security among the DDOS attack.**

***Index Terms*— Detection rate, extreme learning machine, random forest, support vector machine.**

I. INTRODUCTION

Intrusion is a prime problem of security breach, because a single instance of intrusion can steal or delete data from computer and network systems in a few seconds. Intrusion can also damage system hardware. Moreover, intrusion can cause huge losses financially and compromise the IT critical infrastructure, thereby leading to information inferiority in cyber war. Consequently, intrusion detection is important and its prevention is necessary. Different intrusion detection techniques are available, but their accuracy remains an issue; accuracy depends on detection and false alarm rate. The problem on accuracy needs to be addressed to reduce the false alarms rate and to increase the detection rate. This notion was the impetus for this research work. Thus, support vector machine (SVM), random forest (RF), and extreme learning machine (ELM)

are applied in this work; these methods have been proven effective in their capability to address the classification problem. Intrusion detection mechanisms are validated on a standard dataset, KDD. This work used the NSL–knowledge discovery and data mining (KDD) dataset, which is an improved form of the KDD and is considered a benchmark in the evaluation of intrusion detection methods.

II. EXISTING METHOLOGY

In Mobile ad-hoc networks devices are under a capability of wireless communications and networking which makes them able to communicate with each other without the aid of any centralized system. This is an individual system in which nodes are connected by wireless links and send data to each other As we know that there is no any centralized system so routing is done by node itself. According to its mobility and self routing capability nature, there are many weaknesses in its security. One of the major attacks to be assumed in ad hoc network is DDoS attack A DDoS attack is launched by sending huge amount of packets to the target node through the co-ordination of large amount of hosts which are distributed all over in the network. At the victim side this large traffic saves the bandwidth and not allows any other important packet reached to the victim.

III. PROPOSED SYSTEM

We proposed to solve the security issues we need an intrusion detection system. This can be categorized into two models:

1. Signature-based intrusion detection
2. Anomaly-based intrusion detection

The benefits of this IDS technique are that it can be able to detect attack without prior knowledge of attack. Intrusion attack is very simple in wireless network as compare to wired network. One of the major attacks to be considered in ad hoc network is DDoS attack.

A. BLOCK DIAGRAM

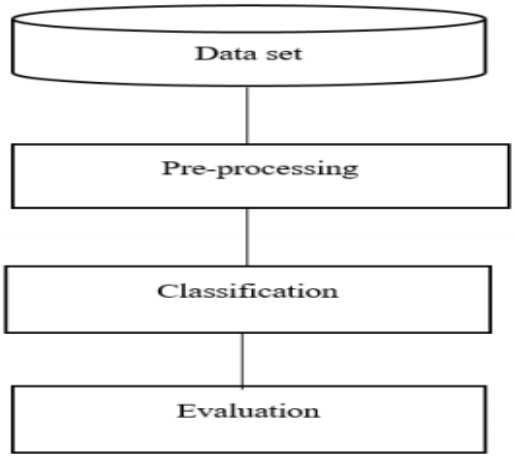


Fig.1. Block Diagram

B. METHODOLOGY

Placing an activity into ordinary and intrusive categories is the core function of an intrusion detection system, which is known as an intrusive analysis engine. Thus, different classifiers have been applied as intrusive analysis engines in intrusion detection in the literature, such as multilayer perceptron, SVM, naive Bayes.

C. Support Vector Machine

SVMs were initially proposed by Vapnik (1995) for solving problems of classification and regression analysis. SWM is a supervised learning technique that is trained to classify different categories of data from various disciplines. These have been used for two-class classification problems and are applicable on both linear and non-linear data classification tasks. SVM creates a hyper plane or multiple hyperplanes in a high-dimensional space, and the best hyperplane in them is the one that optimally divides data into different classes with the largest separation between the classes. A non-linear classifier uses various kernel functions to estimate the margins. The main objective of these kernel functions (i.e., linear, polynomial, radial basis, and sigmoid) is to maximize margins between hyper-planes. Recently, many highly promising applications have been developed by researchers because of the increasing interest in SVMs.

D. Random Forest

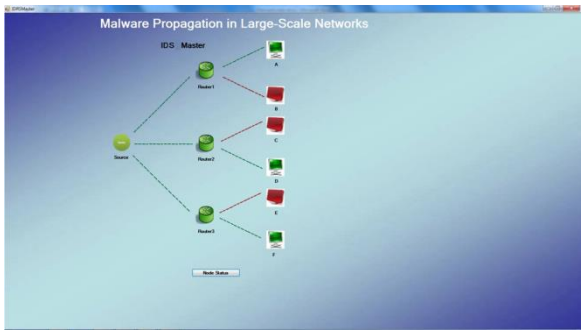
RFs are ensemble classifiers, which are used for classification and regression analysis on the intrusion detection data. RF works by creating various decision trees in the training phase and output class labels those have the majority vote. RF attains high classification accuracy and can handle outliers and noise in the data. RF is used in this work because it is less susceptible to over-fitting and it has previously shown good classification results.

E. Extreme Learning Machine

ELM is another name for single or multiple hidden layer feed forward neural networks. ELM can be used to solve various classification, clustering, regression, and feature engineering problems. This learning algorithm involves input layer, one or multiple hidden layers and the output layer. In the traditional neural networks, the tasks of adjustment of the input and hidden layer weights are very computationally expensive and time-consuming because it requires multiple

rounds to converge. To overcome this problem, Huang et al. proposed an SLFN by arbitrarily selecting input weights and hidden layer biases to minimize the training time. The comprehensive detail of ELM The authors claim that these models learn faster and attain higher generalization capability as compared with other feed forward network models . ELM performance is comparable with SVM or other state-of-the-art machine learning classifiers.

IV. EXPERIMENTAL EVALUATION



V. CONCLUSION

We entirely explore the problem of malware distribution at large-scale networks. The solution to this problem is desperately finish by cyber defenders as the network security community does not yet have solid answers. Different from previous modeling methods, we propose a two layer epidemic model: the upper layer focuses on networks of a large scale networks, for example, domains of the Internet; the lower layer highlights on the hosts of a given network. This two layer model improves the accuracy compared with the available single layer epidemic models in malware modeling. Futherover, the proposed two layer model offers us the distribution of malware in terms of the low layer networks. We execute a restricted analysis based on the proposed model, and obtain three conclusions: The distribution for a given malware in terms of networks follows exponential distribution, power law distribution with a short exponential tail, and power law distribution, at its early, late, and final stage, respectively. In order to examine our theoretical findings, we have conducted extensive experiments based on two real-world large-scale malware, and the results confirm our theoretical claims.

REFERENCES

[1] H. Wang, J. Gu, S. Wang, An effective intrusion detection framework based on SVM with feature augmentation, Knowledge-Based Systems, Volume 136, 2017.

[2] F. Kuang, X. Weihong , S. Zhang, A novel hybrid KPCA and SVM with GA model for intrusion detection, Applied Soft Computing, Volume 18,2014.

[3] A. A. Aburomman, M.B. Reaz, A novel SVM-kNN-PSO ensemble method for intrusion detection system, Applied Soft Computing, Volume 38, 2016.

[4] M.R. Raman, N. Somu, K. Kirthivasan, R. Liscano, V.S. Sriram, An efficient intrusion detection system based on hypergraph - Genetic algorithm for parameter optimization and feature selection in support vector machine, Knowledge-Based Systems, Volume 134, 2017.

[5] S. Teng, N. Wu, H. Zhu, L. Teng and W. Zhang, "SVM-DT-based adaptive and collaborative intrusion detection," in IEEE/CAA Journal of Automatica Sinica, vol. 5, no. 1, pp. 108-118, Jan. 2018.
[6] N.Farnaaz, M.A. Jabbar, Random Forest Modeling for Network Intrusion Detection System, Procedia Computer Science, Volume 89, 2016.