"A Study on Intrusion Detection System for Network and Host"

Monika Chauhan[1] Dr. Shishir Kumar[2]

[1]Research Scholar, Monad University

[2]Professor, JUET Guna

Abstract – Host intrusion detection systems (HIDS) and network intrusion detection systems (NIDS) are methods of security management for computers and networks. In HIDS, anti-threat applications such as firewalls, antivirus software and spyware-detection programs are installed on every network computer that has two-way access to the outside environment such as the Internet. In NIDS, anti-threat software is installed only at specific points such as servers that interface between the outside environment and the network segment to be protected. All methods of intrusion detection (ID) involve the gathering and analysis of information from various areas within a computer or network to identify possible threats posed by hackers and crackers inside or outside the organization. Host-based and network-based ID systems have their respective advantages and limitations. The most effective protection for a proprietary network is provided by a combination of both technologies.

----------♦----------

INTRODUCTION

Intrusion Detection is an emergent trade used against illegal network activities. There is no absolute solution to detect attacks but combinations of different tools and methods do give a quite good result. To give better accuracy in terms of detection, the individual type of IDS is not sufficient. In order to achieve a higher detection rate and low false alarms, more than one type of IDS is required. Like, in this work, NIDS and HIDS are used to detect anomalies using supervised learning. The SVM with GNP classifier is used both in the NIDS and HIDS to classify the data into normal or abnormal.

The proposed system architecture shown in Figure 1 has two major functions, namely, Incremental Support Vector Machine and Genetic Network Programming. The ISVM mainly does the sample selection according to the RBF-kernel function. The GNP rule creator creates the rule to classify the dataset into normal or attack pattern.

The KDD CUP 1999 Dataset for NIDS or Widows Registry for HIDS is taken as the training data and that data set is preprocessed by RBF based kernel function used in ISVM. In case of NIDS, 9 features out of 41 attributes from KDD CUP 1999 Dataset have been extracted for further process. Fuzzification is applied on the training dataset to enhance the rule creation. Genetic Network Programming is applied to the fuzzified dataset and the trained Network Intrusion Detection Model has been produced. The training dataset is used to create the rules and generated signatures are stored for pattern matching. In the testing phase the test dataset has been pattern matched with the NIDS model and finally the data are classified into normal attack or abnormal attack.

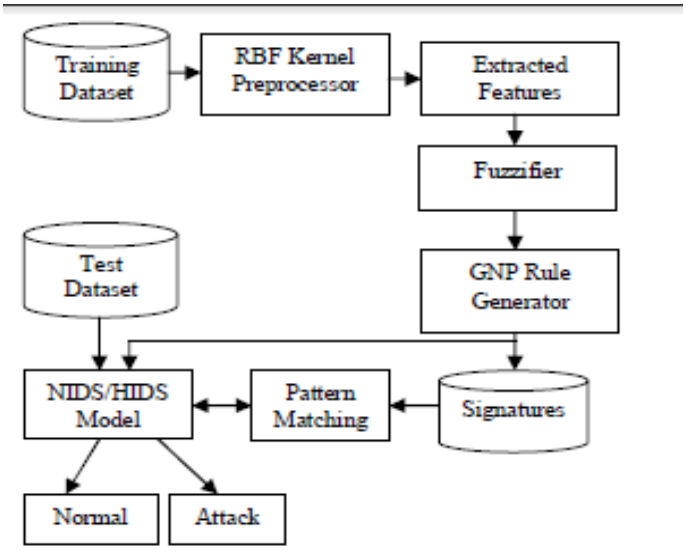

Figure 1: System Architecture for NIDS and HIDS

IMPLEMENTATION DETAILS

In this work, support vector machine is used to preprocess the data using RBF-Kernel function. After preprocessing the data set is classified into the positive kernel and negative kernel. Thus the sample dataset is selected by the kernel function. Fuzzification is applied to the sample dataset and that divides the attributes into sub-attributes which are used for creating the rules easily. Then the genetic network programming is used to create the rules for the intrusion detection in the network. Genetic Network Programming is a step-by-step process which has processing node and judgement node.

Processing nodes are used to pass or just process the rules and the judgement nodes are used to check the rules. If the first rule is satisfied it will then go to next processing node or another judgement node until the rule is created.

PERFORMANCE EVALUATION

This experiment has been done through the 10% of the KDD CUP 1999 Dataset. The dataset is preprocessed by incremental SVM using RBF kernel function and rules are generated using genetic network programming.

The proposed system detects misuse activities with significant improvement in terms of high detection rate and low false positive rate. This can be compared with traditional SVM method and results are shown in Figures 2 and 3.

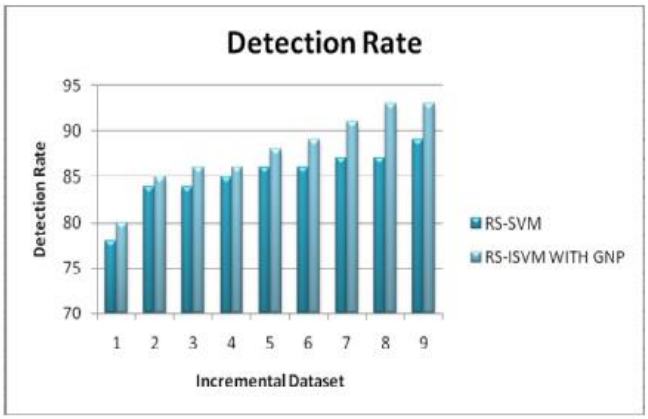

Figure 2 Detection Rate

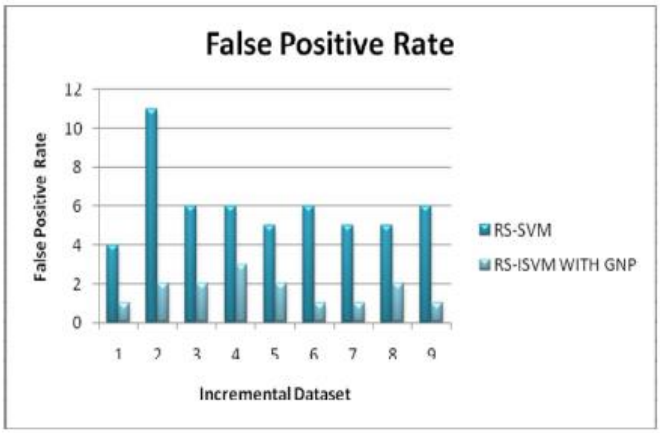

Figure 3 False Positive Rate

From the results shown in the graphs, it can be observed that the proposed RS-ISVM with GNP provided better detection rate and reduced false positive rate when it is compared with the RS-SVM algorithm Intrusion Detection Systems identify intrusions by comparing observable behavior against suspicious patterns. They can be NIDS or HIDS.

The problem with current intrusion detection systems is that existing IDSs are built with either signature-based or anomaly-based systems. When they are used independently, the following problems are being encountered.

1. Signature-based IDS cannot detect unknown attacks without any pre-collected signatures.

2. An anomaly-based system treats any network connection violating the normal profile as an anomaly resulting in frequent reporting of unremarkable data.

3. Traditional IDSs work in isolation and may be easily compromised by unknown or new threats.

A Hybrid IDS technique combines both the advantages of signature based misuse detection schemes and anomaly detection schemes to improve the efficiency of the system. However, the single Hybrid IDS can be easily compromised by other threats. In this work, the Collaborative Hybrid Intrusion Detection System is proposed and deployed for the detection of Denial of Service (DoS) attacks in the network level using Collaborative approach. The collaboration between the hybrid IDSs is designed to enhance the detection of intruders. This collaboration is based on trustworthiness between the peer IDS, calculated using the feedback aggregated from test messages. The intrusion consultation

messages are sent to Hybrid IDS in acquaintance list that is also filled from mature probation list IDS.

The network attacks are increasing day by day with the passing of time. Therefore securing of network resources is important especially from DOS attacks. Denial of Service attacks can essentially disable the computer or the network. Depending on the nature of the enterprise, this can effectively disable the organization. Some DoS attacks can be executed with limited resources against a large, sophisticated site.

PROPOSED SYSTEM

A hybrid system for intrusion detection that combines anomaly detection and misuse detection**.** This involves combining signature based misuse detection system along with anomaly detection scheme thereby eliminating the drawbacks of using these schemes independently.

The isolated IDS can be easily compromised. So to overcome this weakness, the IDS is proposed with CIDN and it has the collective information, knowledge and experience shared by other peers. This enhances the overall accuracy of intrusion assessment as well as the ability of detecting new classes of intrusions.

GNP based Rule Creation

The rule creation is based on genetic network programming. Genetic Network programming is the method to create the rules by passing the attributes to the processing node and judgement nodes. Judgement nodes are the nodes to compare the attributes with already created rules. Processing nodes are the nodes to just pass the attribute from one node to another node.

There are two key points that have been taken into consideration, one is how to select samples to construct the reserved set, and the other is how to assign weights to each sample. The decision-function of ISVM is determined by the support vectors and then determines which samples are most likely to be support vectors.

The KDD CUP 1999 Dataset is too large for the first part of the experiment, so we selected a subset of KDD, about 10% of it, as the input of the proposed system. To avoid the imbalance of the dataset, the sample set are preprocessed to regenerate new training and test set. The preprocessed dataset is produced by the ISVM method. The training time increases according to the size of the dataset increases.

In the training dataset 9 attributes are taken instead of 41 attributes which reduce the noise as well as the training time. The selected attributes are server count, count, source bytes, destination bytes, diff_host_srv_rate, dst_host_port_rate, service, flag, proto_type. In the testing phase, the dataset is compared with the training data model and produce the result according to the type of the attack. The attacks are classified based on the rules generated using the combination of attributes like service, flag, source bytes and destination bytes.

CONCLUSION

IDS supports verbose logging, many events are logged in days, ensure that only pertinent data is collected and that you do not get inundated with unnecessary data. HIDS has more logging than NIDS when taking into account that HIDS logs all machines on the network this is not surprising. If one is looking at HIDS or NIDS it will be good to find a vendor that has good technical backup and that has the pattern files streaming out when there are new vulnerabilities released into the wild much like an antivirus application. If there is LAN bandwidth constraints, it is very feasible to look at a HIDS. If price is an issue then some NIDS solutions are considerably more expensive when compared to a HIDS solution as there is a capital outlay on the hardware and some vendors charge considerably more for the software.

REFERENCES

- Chen, H., Wang, Q. and Shen, Y. “Decision Tree Support Vector Machine Based on Genetic Algorithm for Multi-class Classification”, in Journal of Systems Engineering and Electronics, Vol. 22, No. 2, pp. 322-326, 2011.
- Chirillo, J. “Network Security for Windows, UNIX and Linux Networks: Hack Attacks Denied”, Wiley Publishing Inc., 2nd Edition, 2002.
- Chou, T. and Chou, T. “Hybrid Classifier Systems for Intrusion Detection”, in 7th Annual Communication Networks and Services Research Conference, pp. 286-291, 2009.
- Chowdhury, N. and Murthy, C.A. “Minimum Spanning Tree Based Clustering Techniques: Relationship with Bayes Classifier”, Pattern Recognition, Vol. 30, No. 11, pp. 1919-1929, 1997.
- Cohen, W. “Fast Effective Rule Induction”, in Proceedings of 12th International Conference on Machine Learning, pp. 115-123, 1995.
- Cortes, C. and Vapnik, V. “Support-Vector Networks", Journal on Machine Learning, Vol. 20, No. 3, pp. 273-297, 1995.

- Dartigue, C., Jang, H.I. and Zeng, W. "A New Data-Mining Based Approach for Network Intrusion Detection," in Proceedings of the 7th Annual Communication Networks and Services Research Conference, IEEE Computer Society, pp. 372-377, 2009.
- Dash, M. and Liu, H. "Feature Selection for Classification", in International Journal of Intelligent Data Analysis, Vol. 1, No. 3,pp. 131-156, 1997.
- Dash, M. and Liu, H. "Consistency-Based Search in Feature Selection", Journal on Artificial Intelligence, Vol. 151, No. 1-2, pp. 155-176, 2003.
- Debar H., Dacier M. and Wespi A. "Towards a Taxonomy of Intrusion-Detection Systems", Computer Networks, pp. 805-822, 1992.
- Denning, D.E. "An Intrusion-Detection Model", in IEEE Transactions on Software Engineering, Vol.13, No. 2, pp. 222-232, 1987.