“Hybrid Intrusion Detection System and Its Architecture”

Vikas Verma

Research Scholar of NIMS University, Gujarat

Abstract – The hybrid IDS is obtained by combining packet header anomaly detection (PHAD) and network traffic anomaly detection (NETAD) which are anomaly-based IDSs with the misuse-based IDS Snort which is an open-source project. Intrusions detection systems (IDSs) are systems that try to detect attacks as they occur or after the attacks took place. IDSs collect network traffic information from some point on the network or computer system and then use this information to secure the network. Intrusion detection systems can be misuse-detection or anomaly detection based. Misuse-detection based IDSs can only detect known attacks whereas anomaly detection based IDSs can also detect new attacks by using heuristic methods. In this paper we propose a hybrid IDS by combining the two approaches in one system.

---♦---

INTRODUCTION

The proposed hybrid IDS with collaboration is effective in terms of detection accuracy and detection time. This proposed hybrid system consists of two major subsystems namely the Data Preparation and Collaborative Hybrid Intrusion Detection System. The subsystems in this hybrid system consist of subcomponents such as Feature Extractor, Misuse Detection System (MDS), Anomaly Detection System and Collaborative Intrusion Detection Network. The overall system architecture of the Hybrid Intrusion Detection System proposed in this research work is shown in Figure 1.

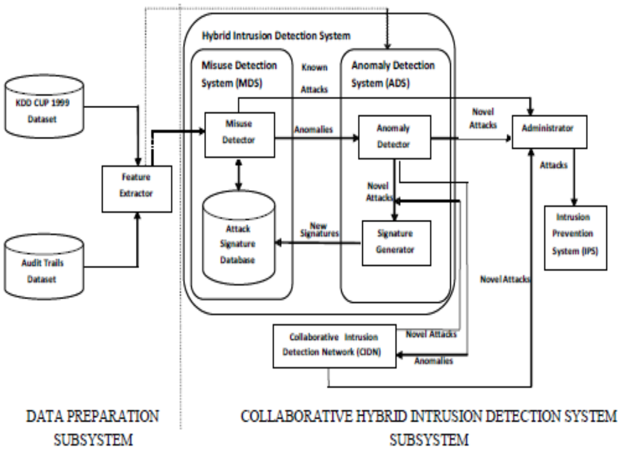
Hybrid Intrusion Detection System
Misuse Detection System (MDS)
Known Attacks
Anomaly Detection System (ADS)
KDD CUP 1999 Dataset
Misuse Detector
Anomalies
Anomaly Detector
Novel Attacks
Administrator
Feature Extractor
Attacks
Novel Attacks
Attack Signature Database
New Signatures
Signature Generator
Intrusion Prevention System (IPS)
Audit Trails Dataset
Novel Attacks
Collaborative Intrusion Detection Network (CIDN)
Novel Attacks
Anomalies
DATA PREPARATION SUBSYSTEM
COLLABORATIVE HYBRID INTRUSION DETECTION SYSTEM SUBSYSTEM

Figure 1: Overall System Architecture

This subsystem is used for effective data selection to enhance the performance of the intrusion detection system. It contains the feature extractor component to extract the required features from KDD CUP 1999 Dataset for the process of intrusion detection. The dataset has been formed out of the selected features and the reduced instances of the KDD CUP 1999 Dataset for classification. This dataset is divided into two sub-datasets. In one sub-dataset, class labels have been removed and are given as input to unsupervised anomaly intrusion detection. The second sub-dataset is supplied as input to the trainer in the supervised misuse intrusion detection, which has the class label. Audit Trails Dataset is taken as an input to the host-based hybrid intrusion detection system. In this, user profile has been considered and set of user behavior (behavior-set) is given to this detection system.

The Feature Extractor subcomponent extracts the relevant features (feature selection) from the database using efficient data pre-processing techniques namely Genetic Algorithm, Incremental Support Vector Machine and Discriminant Analysis according to the type of intrusion detection used in the system. If the features extracted are carefully chosen, it is expected that the feature set extracts the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input. Then, the refined dataset is classified into training dataset and test dataset. Training dataset is used to train the system by using effective classifiers and test dataset is given to the classifiers for classifying them as normal and abnormal patterns.

REVIEW OF LITERATURE

Hybrid intrusion detection systems combine the benefits of signature-based and anomaly-based intrusion detection mechanisms. The combination of network-based and host-based also constitutes the hybrid intrusion detection system. Many such hybrid intrusion detection systems using data mining techniques (Fan et al 2001, Botha et al 2002) are available and discussed below.

Another approach used for classifying the data in data mining is 3-level classifier. Xiang and Lim (2002) proposed a multiple-level hybrid classifier for intrusion detection that uses a combination of Decision tree-classifier and Bayesian clustering algorithm to detect intrusions. The performance of this new algorithm was compared by the authors with other popular approaches and found that significant improvement has been achieved by them from the viewpoint of both high intrusion detection rate and reasonably low false alarm rate.

Behavior based intrusion detection system can detect only those intrusions which cause host behavior to change significantly. Bashah et al (2005) introduces a hybrid system with a technique implemented in conjunction with a behavior based intrusion detection system and misuse detection system. Such a hybrid system would make a “complete intrusion detection system”. The behavior-set can also be refined based on the environment in which the intrusion detection system is used. They have introduced a hybrid intelligent intrusion detection system which was developed by integrating anomaly, misuse and host based detection. A fuzzy logic with network profiling is used to detect the attacks in network data. For host based intrusion detection, neural-networks along with self-organizing maps is used.

Weon et al (2005) proposed a hybrid intrusion detection system which actually reduces the false alarm to 60% level with Snort and retain most of Snort true alarms. The Snort is capable of finding new attacks that cannot be detected by other systems. While many other systems have tried data mining techniques to reduce false alarms, these authors have shown that if a rich set of analyzed data are available, memory based supervised learning can be used to improve the performance of signature based IDS. The experiments carried out by them have shown that the reduction rate is improved significantly.

According to Depren et al (2005), IIDS is a dynamic defensive system that is capable of adapting to dynamically changing traffic pattern and is present throughout the network rather than only at its boundaries, thus helping to catch all types of attacks. The hybrid intrusion detection system using NN and clustering to categorize program behavior as normal or intrusive action has been presented by Zheng et al (2005).

In their work, the entire audit data were first divided into subspaces using the K-means clustering algorithm. After, a set of NNs is used to learn each subspace for intrusion detection separately. During training, NN could recognize normal and abnormal behaviors quickly because audit data, which are in the same subspace, have the similar behavior characters.

The main theme of the intrusion detection is classifying the data according to the similar properties. The k-means and k-medoid algorithm were mainly used in the past to classify the data. Krishnamoorthi et al (2006) proposed a hybrid IDS combines the rule based classifier with simple k-means clustering algorithm to detect intrusion. The initial prototypes developed by JRip classifier (RIPPER) with different granularity of the best rule set were used to test the performance of the hybrid model. The overall classification accuracy offered by hybrid model for the test data set is quite good and comparable with JRip classifier and k-means clustering with random initial prototype.

The hybrid system combines the advantages of low false-positive rate of signature-based IDS and the ability of an Anomaly Detection System (ADS) to detect novel unknown attacks. By mining anomalous traffic episodes from Internet connections, Hwang et al (2007), build an ADS that detects anomalies beyond the capabilities of signature-based SNORT or Bro systems. A weighted signature generation scheme is developed to integrate ADS with SNORT by extracting signatures from anomalies detected. Hybrid IDS extracts signatures from the output of ADS and adds them into the SNORT signature database for fast and accurate intrusion detection. The Hybrid IDS approach proves the vitality of detecting intrusions and anomalies, simultaneously, by automated data mining and signature generation over Internet connection episodes.

Currently, there is an increasing interest to augment fuzzy systems with learning and adaptation capabilities. Neural fuzzy systems and genetic fuzzy systems hybridize the approximate reasoning method of fuzzy systems with the learning capabilities of neural networks and evolutionary algorithms.

Bridges and Vaughn (2000) developed a method that uses GA to detect both network misuses and anomalies. In most of the existing GA based IDSs, the quantitative features of network audit data are either ignored or simply treated, though such features are often involved in intrusion detection. This is because of the large cardinalities of quantitative features. They proposed a way to include quantitative features by introducing fuzzy numerical functions.

COLLABORATIVE HYBRID INTRUSION DETECTION SYSTEM

This subsystem is employed for detecting the unseen intrusive patterns from the network traffic data or audit trail data. It contains the subcomponents namely Misuse Detection System, Anomaly Detection System, Collaborative Intrusion Detection Network and Administrator. The misused detection system and anomaly detection system constitutes a hybrid intrusion detection system. The identified known and novel attacks are directed to the Administrator who raises the alarm as well as takes the appropriate actions for preventing the system.

Misuse or Signature-based Detection System

Misuse Detection System or Signature-based IDS employs a priori knowledge of attack signatures. The signatures are manually constructed by security experts analyzing previous attacks. The collected signatures in the attack signature database are used to match with incoming traffic to detect intrusions. Whenever incoming traffic goes out of the normal profile, anomalous system behavior is identified. In this research work, the effective classifiers are used to classify the data as normal or unseen attacks. Such classifiers used in this misuse detection system of the hybrid intrusion system are Reserved Set-Incremental Support Vector Machine with Genetic Network Programming, Fuzzy Class-Association Rule Mining and Pattern Matching for the source data of network or host.

Anomaly Detection System

Anomaly Detection System treats any network connection violating the normal profile as an anomaly. A network anomaly is revealed if the incoming traffic pattern deviates from the normal profiles significantly. Through a data mining approach, anomaly detection discovers temporal characteristics of network traffic. This system can detect unknown attacks and handles multi-connection attacks well. In this research work, the supervised and unsupervised algorithms are used to detect unseen or novel attacks. The signatures of such detected attacks are generated and stored in the attack signature database for future detection which will act as misuse detection. The data mining algorithms for anomaly detection such as an Enhanced Self-Organizing Map, Minimum Spanning Tree-based Genetic Clustering and Canopy and K-Means Clustering have been implemented to detect unseen attacks.

Collaborative Intrusion Detection Network

The collaboration framework is built around the IDS component. Collaborative Intrusion Detection Network framework connects Hybrid IDSs to form a collaborative network, in which each Hybrid IDS is free to choose the node or network with which to collaborate. When a Hybrid IDS detects suspicious behavior but lacks expertise to make a decision whether it should raise an alarm, it may send requests to its acquainted Hybrid IDSs for consultation. Feedback from the acquaintances is aggregated and a final decision can be made based on the aggregated results. The alert information provided to acquaintances depends on the trust level of each acquaintance. In this research work, a Dirichlet-based trust management algorithm is implemented to achieve Collaborative Intrusion Detection Network which actively detects the anomalies and confirmed attacks are reported to the administrator.

Administrator

The Administrator receives outputs from the above detection systems and reports them to the user through different means like Graphical User Interface (GUI), log files or email. The significant role of the administrator is to react to the detected intrusions in order to prevent future damage. The active responses like dropping the connectivity of the potential attacker or even counter-attacks. A response may be triggered automatically or manually via the user interface. This kind of response will be taken care by Intrusion Prevention System.

The key contributions of this research work with respect to the system architecture are the proposal of a hybrid intrusion detection system which comprises a misuse detection system and anomaly detection system with effective collaboration. In this work, the effective data mining techniques have been employed in extracting relevant features as well as in classifying known and unknown attacks from normal patterns. Finally, this proposed system achieves a high detection rate and low false alarm rate in comparison with the existing individual and hybrid IDSs.

CONCLUSION

In this study, the researcher design and implement the model to perform hybrid model with different metrics. Finally, the results demonstrated by the hybrid model compared with other primary algorithms that have better performance in detecting intrusion. The result have high Accuracy to detect intrusion while using hybrid model rather than primary algorithms, also result shows good percentage of alarms in terms of: False positive, True positive, False Negative and True Negative when researcher use hybrid model.

Although it is true that combining multiple different IDS technologies into a single system can theoretically produce a much stronger IDS, these hybrid systems are not always better systems. Different IDS technologies examine traffic and look for intrusive activity in different ways. The major drawback to a hybrid IDS is getting these different technologies to interoperate successfully and efficiently. Getting multiple IDS approaches to coexist in a single system can be a very challenging task.

REFERENCES

- Abadeh, M.S., Habibi, J. and Lucas, C. “Intrusion Detection using a Fuzzy Genetics-based Learning Algorithm”, Journal of Network and Computer Applications, Science Direct, Vol. 30, No. 1, pp. 414-428, 2007.

- Aggarwal, C.C. and Yu, S.P. “Finding Generalized Projected Clusters in High imensional Spaces”, in Proceedings of the ACM SIGMOD Conference, Vol. 29, No. 2, pp. 70-81, 2000.

- Agrawal, R. and Srikant, R. “Fast Algorithms for Mining Association Rules in Large Databases”, in Proceedings of the 20th International Conference on Very Large Databases, Santiago, Chile, pp. 487-499,1994.

- Anderson, J.P. “Computer Security Threat Monitoring and Surveillance”, Technical Report, Fort, Washington, 1980.

- Anderson, D., Lunt, T., Javitz, H., Tamaru, A. and Valdes, A. “Detecting Unusual Program Behavior using the Statistical Component of the Next-Generation Intrusion Detection Expert System (NIDES)", Technical Report, Computer Science Laboratory, SRI International, Menlo Park, CA, May 1995.

- Apap, F., Honig, A., Hershkop, S., Eskin, E. and Stolfo, S. “Detecting Malicious Software by Monitoring Anomalous Windows Registry Accesses”, in 5th International Symposium on Recent Advances in Intrusion Detection, Zurich, Switzerland, pp. 36-53, 2002.

- Asghar, S. and Iqbal, K. “Automated Data Mining Techniques: A Critical Literature Review”, in International Conference on Information Management and Engineering, pp. 75-79, 2009.

- Ashfaq, M.S., Farooq, U. and Karim, A. “Efficient Rule Generation for Cost-Sensitive Misuse Detection using Genetic Algorithms”, in Proceedings of IEEE International Conference on Computational Intelligence and Security, Vol. 1, pp. 282-285, 2006.