World Applied Sciences Journal 29 (Computer Sciences, Engineering and Its Applications): 53-56, 2014
ISSN 1818-4952
© IDOSI Publications, 2014
DOI: 10.5829/idosi.wasj.2014.29.csea.2221

Dynamic Model of Intelligent Intrusion Detection System

R. Udayakumar, K.P. Thooyamani Khanaa and A.V. Allin Geo

School of Computing Science, Bharath University, Chennai – 73, India

Abstract: Information security is an issue of serious global concern. The complexity, accessibility and openness of the Internet have served to increase the security risk of information systems tremendously. Artificial Intelligence plays a driving role in security services. This article concerns intrusion detection and proposes a method for intrusion detection based on specific Artificial Intelligent approach for intrusion detection. The techniques that are being used include neural network with simple data mining technique to process the network data. The key idea is to discover useful patterns or features that describe user behavior on a system and use the set of relevant features to build classifiers that can recognize anomalies and known intrusions.

Key words: Intrusion Detection · Network Security · Neural Network and Data Mining

INTRODUCTION

Information has become an organization's most precious asset. Organizations have become increasingly dependent on it since more information is being stored and processed on network-based systems. The wide spread use of e-commerce, has increased the necessity of protecting the system to a very high extend. Confidentiality, Integrity and availability of information are the three major concerns in the development and exploitation of network based computer systems. Intrusion Detection System, can detect, prevent and react to the attacks. Intrusion Detection has become an integral part of the information security process.

An Overview of Current Intrusion Detection Systems: Intrusion Detection is defined [1] as the process of intelligently monitoring the events occurring in a computer system or network and analyzing them for signs of violations of the security policy. The primary aim of Intrusion Detection Systems (IDS) is to protect the availability, confidentiality and integrity of critical networked information systems. Intrusion Detection Systems (IDS) are defined by both the method used to detect attacks and the placement of the IDS on the network [3]. IDS may perform either misuse detection or anomaly detection and may be deployed as either a network-based system or a host-based system. This result in four general groups: misuse-host, misuse-network, anomaly-host and anomaly-network. Misuse detection relies on matching known patterns of hostile activity against databases of past attacks. They are highly effective at identifying known attack and vulnerabilities, but rather poor in identifying new security threats. Anomaly detection will search for something rare or unusual by applying statistical measures or artificial intelligence to compare current activity against historical knowledge [4]. Common problems with anomaly-based systems are that, they often require extensive training data for artificial learning algorithms and they tend to be more computationally expensive, because several metrics are often maintained and these need to be updated against every systems activity.

Some IDS combine qualities from all these categories (usually implementing both misuse and anomaly detection) and are known as hybrid systems. Artificial Intelligence techniques have been applied both to misuse detection and also for anomaly detection. SRI's intrusion Detection Expert System (IDES) [2] encodes an expert's knowledge of known patterns of attack and system vulnerabilities as if-then rules. Time-based Inductive machine (TIM) for intrusion detection [3] learns sequential patterns. Recently, techniques from data mining area have been used to mine normal patterns from audit data [4,5,6]. Several approaches applying artificial neural networks in the intrusion detection system have

Corresponding Author: R. Udayakumar, School of Computing Science, Bharath University, Chennai – 73, India.

been proposed [7,8,9]. NeGPAlM [10] based on trend analysis, fuzzy logic and neural networks is to minimize and control intrusion. Existing intrusion detection especially commercial intrusion detection systems that must resist intrusion attacks are based on misuse detection approach, which means these systems will only be able to detect known attack types and in most cases they tend to be ineffective due to various reasons like non-availability of attack patterns, time consumption for developing new attack patterns, insufficient attack data, etc.

Computer Attack Categories: DARPA [11] categorizes the attacks into five major types Based on the goals and actions of the attacker.

DoS attacks try to make services provided by or to computer users be restricted or denied. For example, SYN-Flood attack, where the attacker floods the victim host with more TCP connection requests than it can handle, causing the host to be unable to respond even to valid requests. Probe attacks attempt to get information about an existing computer or network configuration [8].

Remote-to-local (R2L) attacks are caused by an attacker who only has remote access rights. These attacks occur when the attacker tries to get local access to a computer or network.

User-to-root (U2R) attacks are performed by an attacker who has rights of user level access and tries to obtain super user access.

Data attacks are performed to gain access to some information to which the attacker is not permitted access. Any R2L andU2Rhasagoalofaccessing the secret files.

Data Mining and Association Rules: Data Mining is the automated extraction of previously unrealized information from large data sources for the purpose of supporting actions. The recent rapid development in data mining has made available a wide variety of algorithms; drawn from the fields of statistics, pattern recognition, machine learning and databases. Specifically, data mining approaches have been proposed [4, 12] and used for anomaly detection. Association rule algorithms find correlations between features or attributes used to describe a data set. The most popular algorithm for mining rules based on two-valued attributes is APRIORI.

Data Set Evaluation: The 1998 version of MIT Lincoln Laboratory – DARPA (Defense Advanced Research Projects Agency) intrusion detection evaluation data was used. A subset of the data that contained the desired attack types and a reasonable number of normal events were selected manually.

Attack Types: There are at least four different known categories of computer attacks including denial of service attacks, user to root attacks, remote to user attacks and probing attacks. The type of attack that was included in the data set for the study: guess attack. The factor for choosing the attack type is the availability of the enough data records [13].

Mining Patterns from Audit Data: In order to construct an accurate base classifier, we need to gather a sufficient amount of training data and identify a set of meaningful features. This task requires insight into the nature of the audit data and can be very difficult without proper tools. We use the term “audit data” to refer to general data streams that have been properly processed for detection purposes.

Association Rules: The goal of mining association rule s is to derive multi-feature (attribute) correlations from a database table. For example, an association rule from the shell command history file (which is a stream of command and their arguments) of a user is trn => rec.humor; [0.3, 0.1],

Which indicates that 30% of the time when user invokes trn, he or she is reading the news in rec.humor and reading this newsgroup accounts for 10% of the activities recorded in his or her command history file. The motivation for applying the association rules algorithm to audit record is that there is evidence that program executions and user activities exhibit frequent correlations among system features. The most popular algorithm for mining rules based on two-valued attributes is APRIORI. Used APRIORI algorithm for mining association rules, which are used as inputs to Neural Network

Implementation: Training and Validation Method: The present study was aimed to solve a two-class problem. Here, two cases are described which can be extended to cases with more attack types. An output layer with two neurons (output states) was used. In this paper a two layer neural network means a neural network with one hidden layer (the input layer is not considered because it acts like a buffer and no processing takes place; however, the output layer is considered).

MATLAB Neural network Toolbox was used for the implementation of the network. Using this one can specify number of layers, number of neurons in each layer and number of training epochs [8]. The implemented neural network has 150 input neurons and two output neurons (equal to the number of classes) and the number of neurons in hidden layer is 80. Competitive learning algorithm is used for training. In competitive learning, the elements of the network compete with each other for the "right " to provide the output associated with an input vector. Only one element is allowed to answer the query and this element simultaneously inhibits all other competitors.

Experimental Results: As a first step, the file containing the data set should be processed using the tcpdump program to generate a simple event log without data portion of the packet. The DARPA project provided two list files for comparing the results of intrusion detection. The first one 'bsm.list' is actually produced by using the Sun's Basic Security Monitoring (BSM) tool. The another is 'tcpdump.list' which is a result of another method. Using the two list files collected the normal and attack patterns. Used Christian Borgelt's Apriori algorithm for mining the association rules from the preprocessed DARPA data set. The minimum support and minimum confidences are 80 and the minimum rule length is 3. The following shows the sample association rules mined from the guess attack data set.

> <- ack win IP (98.7%, 100.0%) IP <- ack win > (98.7%, 100.0%) win <- ack IP > (98.7%, 100.0%) ack <- win IP > (99.7%, 99.0%)

The following shows the sample association rules mined from the normal data set. > <- ack win IP (94.3%, 100.0%).

IP <- ack win > (94.3%, 100.0%) win <- ack IP > (94.3%, 100.0%) ack <- win IP > (96.0%, 98.3%)

Association rules are collected which are used for training the Self Organizing Maps. The rules mined are normalized to numeric values and used for training. The network was trained using training data and tested using test data.

RESULTS AND CONCLUSION

An approach for neural network based intrusion detection system, intended to classify the normal and attack patterns and the type of the attack have been presented in this paper. The system is built to run on historical DARPA 1998 training data sets. A great deal of trouble is to get the neural network to detect all types of attacks simultaneously. However the neural network performed well on individual types of attack namely guess attack. The training and testing is limited because our dataset did not contain many instances of the same attack. The following table shows the percentage of prediction accuracy.

Prediction Accuracy for Training		Prediction Accuracy for Testing	
Normal Pattern	Guess Attack Pattern	Normal Pattern	Guess Attack Pattern
80%	100%	80%	100%

REFERENCES

1. Bace R.G., 0000. Intrusion Detection, Technical Publishing (ISBN 1-57870-185-6).
2. Lunt, T., 1993. "Detecting intruders in computer systems". Conference on Auditing and Computer Technology.
3. Udayakumar, R., V. Khanaa and K.P. Kaliyamurthie, 2013. Performance Analysis of Resilient FTTH Architecture with Protection Mechanism, Indian Journal of Science and Technology, ISSN: 0974-6846, 6(6): 4737-4741.
4. Udayakumar, R., V. Khanaa and K.P. Kaliyamurthie, 2013. Optical Ring Architecture Performance Evaluation using ordinary receiver, Indian Journal of Science and Technology, ISSN: 0974-6846, 6(6): 4742-4747.
5. Teng, H., K. Chen and S. Lu, 1990. "Adaptive real time anomaly detection using inductively generated sequential patters".IEEE computer society symposium on research in security and privacy, California, IEEE Computer Society, pp: 278-84.
6. Lee, S. Stolfo and K. Mok, 1998. "Mining audit data to build data to build intrusion detection models". Fourth international conference on knowledge discovery and data mining, New York, AAAI Press, pp: 66-72.
5. Mukkamala, R., J. Gagnon and S. Jaiodia, 2000. Integrating data mining techniques with intrusion detection methods. Research Advances in Database and Information Systems Security, pp: 33-46.
6. Stolfo, S. and Lee, Chan, 2000. "Data mining-based Intrusion detectors : An overview of the Columbia IDS Project" SIGMOD Record, 30(4).

7. Debar, M. Becker and D. Siboni, 1992. “A neural network component for an intrusion detection system”. IEEE Computer Society Symposium on Research in Computer Security and Privacy, pp: 240-250.
8. Saravanan, T., G. Saritha and R. Udayakumar, 2013. A Robust H-Infinity Two Degree of Freedom Control for Electro Magnetic Suspension System, Middle-East Journal of Scientific Research, ISSN:1990-9233, 18(12): 1827-1831.
9. Tan, K., 1995. “The Application of Neural Networks to UNIX Computer security”. IEEE International conference on Neural Networks, 1: 476-481.
10. Wang, J., Z. Wang and K. Dai, 0000. “A Network intrusion detection system based on ANN”, InfoSecu04, ACM 2004(ISBN1-58113-955-1).
11. Botha, M., R. Solms, K. Perry, E. Loubser and G. Yamoyany, 2002. “The utilization of Artificial Intelligence in a Hybrid Intrusion Detection System”, SAICSIT, pp: 149-155.
12. Thooyamani, K.P., V. Khanaa and R. Udayakumar, 2013. Blue tooth broad casting server, Middle-East Journal of Scientific Research, ISSN:1990-9233, 15(12): 1707-1712.
13. Thooyamani, K.P., V. Khanaa and R. Udayakumar, 2013, Improving Web Information gathering for personalised ontology in user profiles, Middle-East Journal of Scientific Research, ISSN:1990-9233(12), pp: 1675-1679.
14. MIT Lincolon Laboratory, 1999 DARPA intrusion detection evaluation design and procedure, DARPA Technical report, Feb 2001.
15. Dokes, P., I. Ertoz, A. Lazarevic, J. Srivastava and P. Tan, 2002. “ Data Mining for Network Intrusion Detection”, Proceedings of NSF Workshop on Next Generation Data Mining.