Intrusion Detection System Using Data Mining

Ashish R. Varma Pankaj S. Desai P.R. Bhaladhare

Abstract — Intrusion Detection Systems (IDSs) have become an efficient defense tool against network attacks since they allow network administrator to detect policy violations. However, traditional IDs are vulnerable to original and novel malicious attacks. Also, it is very inefficient to analyze from a large amount volume data such as possibility logs. In addition, there are high false positives and false negatives for the common lDSs. Data mining has been popularly recognized as an important way to mine useful information from large volumes of data which is noisy, fuzzy, and random. Thus, how to integrate the data mining techniques into the intrusion detection systems has become a hot topic recently. In this paper, we present the whole techniques of the IDS with data mining approaches in details.

Key Words — Intrusion Detection; Network Security; Data Mining

I. INTRODUCTION

With the rapid development World Wide Web and computers during the past decade, security has become a crucial issue for computer systems. A secure network should provide data confidentiality, data integrity, and data availability. Intrusion is an action that tries to destroy data confidentiality, data integrality, and data availability of network information.

The purposed of Intrusion Detection Systems (IDS) is designed to detect attacks against computer systems over insecure networks by this way that detects attempts by legitimate users to abuse their privileges or to exploit security vulnerabilities for comprising the computers.

Existing IDS systems can be divided into two categories according to the detection approaches: anomaly detection and misuse detection or signature detection. Anomaly detection (also called Behavior detection) is an approach to detect intrusions by first learning the characteristics of normal activity of users. Then the system uses such characteristics to judge whether the user's activity is normal or not. Misuse detection systems are the approach that tries to match user activity to stored signatures of known exploits or attacks. That is to say, such detection system use a priori defined knowledge to check whether the new activity is in that knowledge database. If yes, the IDS considers this activity may be as a possible attack and then blocks it.

On the other hand, According to the resources they monitor, IDS systems are divided into three categories: Network-based IDS, Host-based IDS, and Application-based IDS. Network-based intrusion detection systems monitor the network traffic and use these raw network packet's content to analyze network, transport, and application protocols to identify suspicious activity [2]. Host-based IDS monitors a single machine and audits data traced by the hosting operating system. Typical examples of audited data are system calls, event, resource usage, and logs on Windows NT and syslog in UNIX environments [1]. Application-based IDS can be classified to the host-based IDS. It analyzes the events transpiring within a software application.

II. PROCEDURE FOR PAPER SUBMISSION

In this section, we discuss the architecture of IDS with data mining techniques. The key idea in using data mining-based approaches to build secure models for IDS is to generalize from both known attacks and normal behavior in order to detect unknown attacks [9].

2.1 Data Mining-based Approaches IDS

*1) **Classification:** Classification is used to map a data item into one of several pre-defined classes by inputting a training data set and building a model of the class attribute based on the rest of the attributes.*

Based on this theory, it is easy to be applied the algorithm to the IDS. Concretely, we take intrusion detection as classifications issue that each audit data categories to normal or invasive collection. The main techniques for IDS are to collect various aspects of the intrusion data to generalize known instruction structure. When the needed new audit data come, the IDS engines decide whether it is normal data or intrusion categories according to the known instruction structure model.

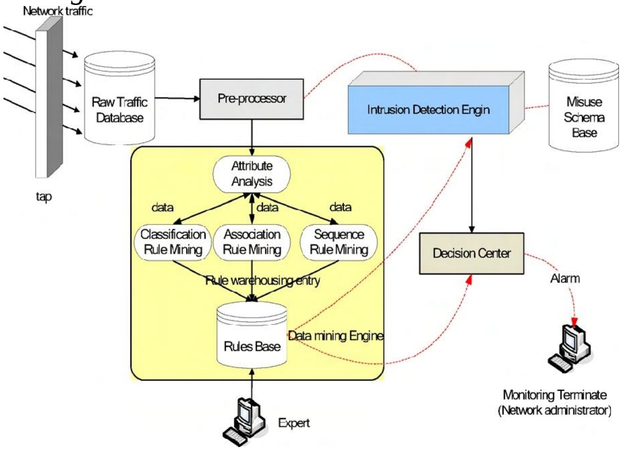
2) Association

Association rules mining is an important subject in the study of data mining. The purposed of association rules mining is to seek to discover associations among transactions encoded in a database. An association rule takes the form of A -> B where A (the antecedent) and B (the consequent) are sets of predicates. Typical mining algorithms are the Apriori and AprioriTid [11]

3) Sequence rules

The purpose of sequence rule mining is to discover a set of attributes, shared access time among a large number of objects in a given database. This algorithm can discover a sequence of events scheme from a security event database. The algorithms focus on the chronological relationship.

A. Figures



CONCLUSION

With the development of the network, new attack approaches will be happen. Thus, excellent IDS not only can detect such new attacks, but also has a low misreport ratio as well. Data mining has been popularly recognized as an important way to mine useful information from large volumes of data which is noisy, fuzzy, and random so that the intrusion detection system based on data mining received much attention recently.

REFERENCES

[1] J.P. Anderson. Computer Security Threat Monitoring and Surveillance. Technical report, James P. Anderson Co., Fort Washington, PA, April 1980.

[2] T.H. Ptacek and T.N. Newsham. lnsertion, Evasion and Denial of Service: Eluding Network lntrusion Detection. Technical report, Secure Networks, January 1998

[3] Fayyad, D, Piatetsky-Shapiro, G., and Smyth, P. From Data Mining to Knowledge Discovery in Databases. Al Magazine, 17(3):37.1 54. 1996a.

[4] Han, J. and Kamber, M. Data Mining: Concepts and Techniques.Morgan Kaufmann Publisher.2000

[5] A. Ghosh and A. Schwartzbard. A study in using neuralnetworks for anomaly and misuse detection. In Proceedingsof the Eighth USENIX Security Symposium, 1999.

[6] L. Pornoy. lntrusion detection with unlabeled data usingclustering. ln Undergraduate Thesis, Columbia University,Department of Computer Science, 2000.

[7] Peng Liu et al .The Design and Implementation ofa Selt Healing Database System. School of lnfo Sciences and Technology Department of Information Systems, PennsylvaniaState University UMBC, University Park, PA 16802 Baltimore,MD 21250.

[8] Pramote Luenam, Peng Liu, .ODAM An On-the-tly DamageAssessment and Repair System for Commercial DatabaseApplications., Dept. of Info. Systems, UMBC Baltimore, MD21250.

[9] Wenke Lee, Salvatore J Stolfo Data Mining Approaches for lntrusion Detection. Proceedings of the 7th USEN1X Security Symposium. 2000

[10] R. Agrawal, T. lmielinski, A.Swami, "Mining association rules between sets of items in large database". In: P. Buneman, S. Jajodia eds. Proc. of 1993 ACM SlGMOD Conf on Management of Data. Washington DC ACM Press, 1993. pp. 207-216

AUTHOR’S PROFILE

Ashish R. Varma,
BE.Computer Science & Engineering.
Pursuing M.tech in Computer Engineering.IET.Alwar(RTU).

Pankaj S. Desai
B.E. Information Technology , Pursuing M.E. in Computer Sci. & Engg . ,GHRIEM , Jalgoan(NMU).

Prof.P.R. Bhaladhare
M.tech Computer(AP & HEAD SNJBCOE,Chandwad.)