International Journal of Security and Its Applications
Vol. 10, No. 8 (2016), pp.145-150
http://dx.doi.org/10.14257/ijsia.2016.10.8.12

"Soft-Man" and Data Mining based Distributed Intrusion Detection System

Jun Zheng

Network Information Center, Baotou Teachers' College
Science Road No. 3, Qingshan District, Baotou, Inner Mongolia, China, 014030
E-mail:zhengjun_025@163.com

Abstract

As modern computer networks are large-scale with numerous nodes, the conventional concentrated intrusion detection system fails to work effectively. To deal with the above situation, the paper proposed a "Soft-Man" and data mining based distributed intrusion detection system (SMDMDIDS, for short). Specifically, it designed an overall structure model of the detection system, expounded the system's communication models, and designed the communication models and cooperation methods between Soft-Mans as well. The paper also defined hierarchical cooperation models for the Soft-Mans and designed corresponding data mining models. Finally, with the help of IDS Informer tools, the paper conducted a simulation experiment on network intrusion detection. The experimental results showed that the proposed intrusion detection system in the paper had good detection performance.

***Keywords**: Soft-Man; Intrusion detection; Data mining*

1. Introduction

Both the scales and network nodes of modern computer networks have been increasing. Under this circumstance, if the conventional intrusion detection model is applied, data processing for all the nodes will be done on the system processing center. This will bring heavy data-processing workload to the system processing center, and cause troubles of serious network blockage, prolonged network delay, and poor performance. Soft-Mans are sent to different network nodes for data collection, data analysis, and command responses. Meanwhile, the central Soft-Man is responsible for an integrated in-depth analysis of the analysis results of the monitoring nodes. The system processing center is merely in charge of intrusion statistics and management components. Since piles of data computation work are distributed to monitoring nodes, there is no more overload for the system processing center, which enhances the real-time performance of the system. In addition, with the constantly enlarging computer network, network attack modes are emerging endlessly. The number of relative data grows rapidly in an astounding manner. It is difficult to analyze masses of audit data with the help of tradition methods, or else the underreporting rate will be high. Applying data-mining technologies to the distributed intrusion detection system can help improve precision and intelligence of system detection.

Therefore, in face of the adverse conditions that the detection host of the concentrated intrusion detection system has already become a safety bottleneck of the network and that safety data of the network has rocketed, the paper proposed SMDMDIDS on the basis of a full advantage of the flexible autonomy of Soft-Man.

ISSN: 1738-9976 IJSIA

2. System Design

The overall structure of SMDMDIDS is shown in Figure 1.

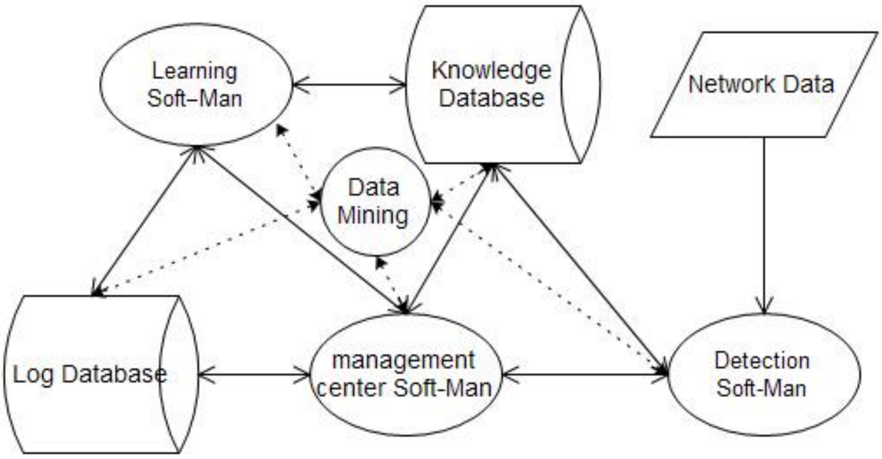

Figure 1. The Overall Structure of SMDMDIDS

In the above model, the knowledge database stores the necessary rules and data of safety strategy for the detection Soft-Man to operate well. The log database stores some fundamental data, including system logs and operating records. The detection Soft-Mans reside in the detection host and operate autonomously, and collect specific network data package together with Snort. They also detect and process some basic abnormal behaviors.

The management center Soft-Mans are composed of tracking Soft-Mans, status detection Soft-Mans, and central management Soft-Mans. The tracking Soft-Mans track routes and sources of intrusions. The status detection Soft-Mans are responsible for detecting data of the detection system and collecting Soft-Mans status, aiming at ensuring smooth communication of the system and system stability. The central management Soft-Mans are in charge of all the Soft-Mans in the management system, the knowledge database, and the log database of the system as well.

The learning Soft-Mans reflect intelligentization of SMDMDIDS. They can learn with the help of data mining algorithm without interfering system detection, and also update working models or establish new models of the knowledge database timely on the basis of learning outcomes. Upon detecting changes in the knowledge database, they will inform the management center Soft-Mans to address all the detection Soft-Mans that the knowledge database has been updated. After the management center Soft-Mans perform the request, the detection Soft-Mans extract new rules and models from the knowledge database.

3. Key Technologies of the System

3.1. The Communication Technologies of SMDMDIDS

In SMDMDIDS, Soft-Mans communicate with each other by means of messages. There are several send modes of system messages as follows:

1. Synchronous messages. This mode is synchronous. The execution of current operations will be suspended until the messages have already been handled.

2. Asynchronous messages. This mode is asynchronous. The execution of current operations continues as the messages are being processed.

3. Broadcast messages. This mode is asynchronous. The messages can be received by the same Soft-Mans.

4. Forwarding messages. For such messages, the targeted messages are forwarded and transferred to other Soft-Mans as parameters.

3.2. Communication Models of SMDMDIDS

In SMDMDIDS, there is a data collection Soft-Man (DCSM, for short) and a Snort in each monitoring node. The DCSM is responsible for analyzing and processing the network data that are collected by the Snort in the same monitoring node. They communicate with each other by means of messages. The corresponding communication model is shown as follows.

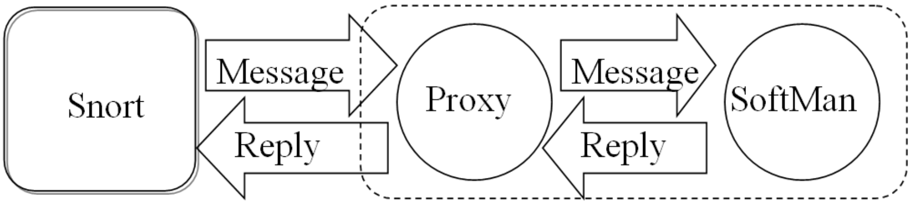

Figure 2. The Communication Model between a Snort and a Soft-Man

There are also data analysis Soft-Mans (DASM) in SMDMDIDS, who cooperate with each other. The status detection Soft-Mans (SDSM) are responsible for detecting the status of DCSM and DASM of the detection system, and are supposed to know the latest conditions of themselves in time by exchanging messages. The central management Soft-Man (CMSM) manages and communicates with all the Soft-Mans of the system.

3.3. The Cooperation Modes of Soft-Mans in SMDMDIDS

The following are several cooperation modes of Soft-Mans in SMDMDIDS.

1) The analytical mode. In each monitoring node of SMDMDIDS has a DCSM responsible for collecting, analyzing and processing log and audit data of the monitoring node. DASM further analyzes the analysis results of DCSM. The relationship between DASM and DCSM is that DCSM is analyzed by DASM.

2) The complementary mode. If two DCSM (DCSM1 and DCSM2, for example) have detected correlated network attacks in different network nodes, the DASM may undertake meta-analysis on both the detection result of DCSM1 and that of DCSM2. In this way, composite network attacks can be detected effectively.

3) The verification mode. If DCSM1 detects a network attack in a network node, it will report the attack data to DASM. After analysis, DASM moves to DCSM2 to observe whether DCSM2 detects the same attack. If so, the network attack for DCSM1 is verified.

4) The control mode. In SMDMDIDS, CMSM manages such functions as the initialization and closure of Soft-Mans, the update of feature base, and the adjustment of detection algorithm in the control system. CMSM cooperates with other Soft-Mans by means of control.

3.4. The Hierarchical Cooperation Model of Soft-Mans in SMDMDIDS

The following is the structure of the Soft-Mans cooperation system.

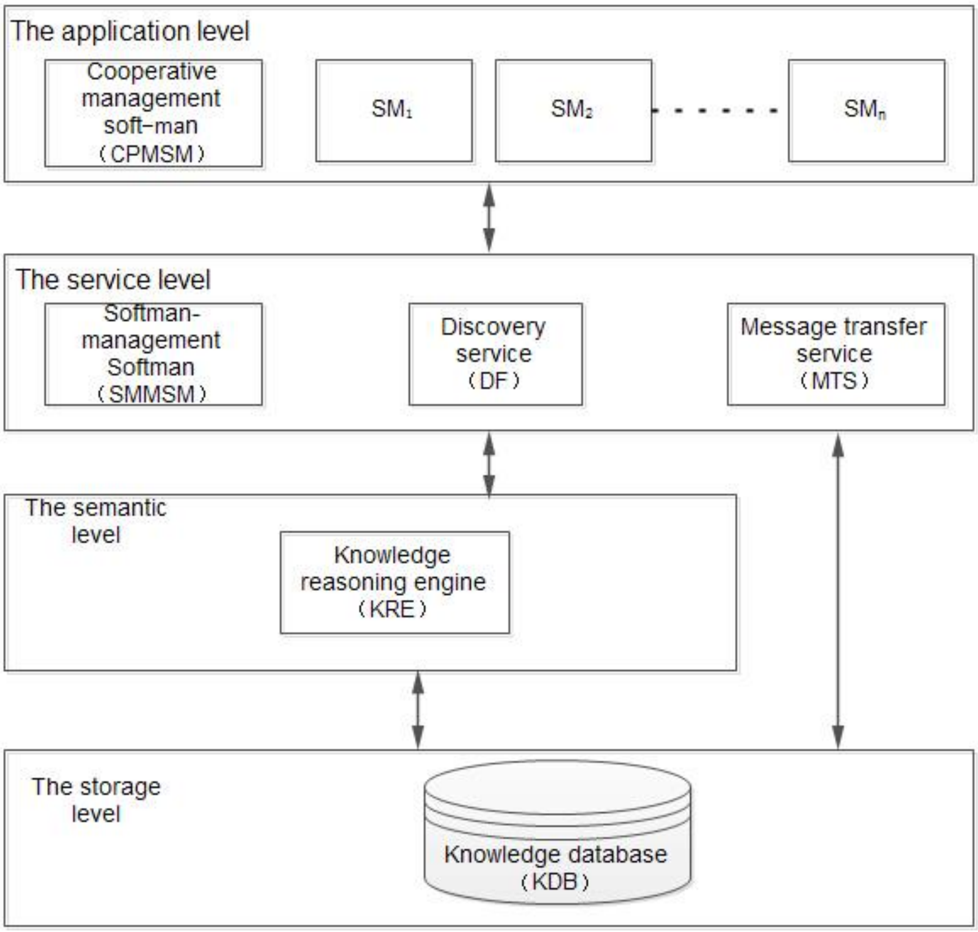

Figure 3. The Structure of the Softmen Cooperation System

As seen from Figure 3, there are four levels in the structure of the Soft-Mans cooperation system: the application level, the service level, the semantic level, and the storage level.

The application level contains cooperative management Soft-Mans (CPRSM, for short) and network Soft-Mans with other functions. CPRSM is designed to register incoming foreign Soft-Mans and log off outgoing foreign Soft-Mans in terms of the cooperation space.

The service level contains Soft-Mans-management Soft-Mans (SMMSM, for short), discovery service (DF, for short), and message transfer service (MTS, for short). SMMSM is responsible for managing and monitoring the cooperation process of other Soft-Mans as well as the knowledge database (KDB, for short). DF assists Soft-Mans to find suitable cooperative Soft-Mans. MTS is used to guarantee the exchange and transfer of messages in the process of mutual cooperation between Soft-Mans.

The semantic level only contains the knowledge reasoning engine (KRE, for short), which provides knowledge reasoning interfaces for KRE and knowledge check interfaces for system knowledge.

The storage level only contains knowledge database (KDB, for short). KDB describes multi-aspect knowledge of intrusion detection, and serves as the foundation of intrusion detection by Soft-Mans. In addition, KDB also describes the cooperation process and modes between Soft-Mans.

3.5. Data Mining Design

The data mining technology offers a decision-supporting process for SMDMDIDS. The data mining design mainly works to explore the rules of both normal behaviors and abnormal behaviors from training data to establish rule base. For data sources with different natures, different data mining algorithms should be adopted so that the implied rules can be discovered. There are mainly four data mining methods in SMDMDIDS.

1) The method of correlation analysis. It aims to excavate the mutual relationship between data. In this method, a group of items and a record set are given. Through an analysis of the record set, the relevance between items can be deducted. This method is used to discover relevant characteristics between various intrusions.

2) Sequence mode analysis. It aims to explore the connection of intrusion orders. Many intrusions by hackers have sequences, and some intrusions is sure to occur after other ones. The method of sequence mode analysis can be used to discover the connection of intrusion orders, and further extract the time series characteristics of intrusions.

3) Classification analysis. It is applied to check previous intrusions and to set classification standards of intrusion modes as well. The classification standards are referred to in classifying and describing each hazard level.

4) Cluster analysis. It aims to divide record sets reasonably according to certain rules, and to describe different grades of record sets explicitly or implicitly. Based on the descriptions in 3), the cluster analysis method is used to redivide users' behavior data, and to describe different divisions of record sets explicitly or implicitly for better results.

4. Analysis and Evaluation of the Experimental Results

IDS Informer v4.0, a simulative network attack software, is applied to SMDMDIDS for simulation tests. Developed by Blade Corporation, IDS Informer is able to generate as many as 600 kinds of attack data that are used for effective intrusion tests. Network data packages of this software are the ones that can record the whole and precise real attacks. After encapsulation and automatic processing, the data packages form a document of attack data. In this way, a grand attack database is built up for detection of various attack identification strategies of SMDMDIDS.

IDS Informer is used for the test in the paper to simulate four kinds of intrusion attacks: PROBE, R2L, U2R and DOS. Each of the kind contains 100,000 normal data and 5,000 intrusion records. The detection results are shown in the following table.

Table 1. The Detection Results

	PROBE	*R2L*	*U2R*	*DOS*
Detection rate	*100*	*98*	*99*	*100*
Error rate	*1.5*	*0.2*	*1.43*	*0*

As can be seen from the results, SMDMDIDS achieves high detection rate in each of the four kinds of intrusion detection. The results are ideal.

Acknowledgements

This work was financially supported by "the College Science Research Project of Inner Mongolia, China "(Grant No. NJZY16238).The author would like to acknowledge and thank for the support.

References

[1] T. Xu-yan, “Generalized Model, Intelligent Simulation, Soft－Man”, Computer Simulation, vol. 7, **(2011)**, pp. 224-228.

[2] E. Xiao-zheng and C. Ding-fang, “Application Research of Aglet Message System”, Journal of Hubei Polytechnic University, vol. 3, **(2004)**, pp. 19-22.

[3] L. Ping-shan and X. Qian-xia, “A Research on the Mechanism of Aglet Programming”, Journal of Guilin Institute of Electronic Technology, vol. 3, **(2002)**, pp. 23-27.

[4] Wiederhold, “Mediators in the Architeeture of Future Information Systems [J]”, IEEE Computer, vol. 25, no. 3, **(1992)**, pp. 38-49.

[5] H. Shan-li and S. Chun-yi, “An Intention Model for Agent”, JOURNAL OF SOFTWARE, vol. 7, **(2000)**, pp. 965-970.

[6] J. Wen-Pin and S. Zhong-Zhi, “A Study of Cooperation Processes in Multi-Agent Systems”, Journal of Computer Research and Development, vol. 8, **(2000)**, pp. 904-911.

[7] Q. Si-han, J. Jian-chun, M. Heng-tai, W. Wei-ping and L. Xue-fei, “Research on intrusion detection techniques: a survey”, Journal of China Institute of Communications, vol. 7, **(2004)**, pp. 19-29.

[8] V. Giovanni, Kemmerer AR and Blix per, “Designing a Web of Highly-Configurable Intrusion Detection Sensors”, RAID, **(2001)**.

[9] L. Tao-shen and T. Ren-peng, “Improved intrusion detection approach based on sequence analysis of system calls”, Computer Engineering and Design, vol. 10, **(2006)**, pp. 1761-1763+1766.

[10] J. Jian-chun, M. Heng-tai, R. Dang-en and Q. Si-han, “A Survey of Intrusion Detection Research on Network Security”, Journal of Software, vol. 11, **(2000)**, pp. 1460-1466.

[11] Jou F., Gong F., Sargor C., Wu X., Wu S. F., Chang H. C. and Wang F., “Design and Implementation of a Scalable Intrusion Detection System for the protection of Network Infrastucture”, IEEE, **(1999)**.

[12] D. Mian, S. gang, H. Qiang and X. Wei, “Modeling and Simulation Research of Active Heave Compensation System. Review of Computer Engineering Studies”, vol. 1, no. 2, **(2014)**, pp. 15-18.

[13] Y. Yan, “A Practice Guide of Predicting Resource Consumption in a Web Server”, Review of Computer Engineering Studies, vol. 2, no. 3, **(2015)**, pp. 1-8.

Authors

Zheng Jun, He is an associate professor of computer science in the School of Baotou Teachers’ College. He received his Master degree in computer science in 2009. His research interests focus on network security, artificial intelligence, machine learning and natural language processing.