Intrusion detection system evaluation model based on model checking

Pengtao XU[1, a], Weijun ZHU[2, b]

[1] Information Engineering, Zhengzhou University, Zhengzhou, 450001, China

[2] Information Engineering, Zhengzhou University, Zhengzhou, 450001, China

[a]email: ieptxu@163.com, [b]email: iewjzhu@zzu.edu.cn

Keywords: Model Checking; Intrusion Detection; Check Data Set; Temporal Logic

Abstract. At present，with the increase of network bandwidth and network data volume, pattern matching to detect intrusion detection systems have many problems to solve. Then, the method based on the model checking method was proposed, and applied to the intrusion detection system[1][2][3][4].However, the Intrusion detection algorithm based on model checking can not meet about comprehensive performance evaluation. Through the analysis of KDDCUP99 data set, combined with the existing logic formula of attack type based on the behavior of the data set.

Introduction

Intrusion Detection System (IDS) can actively find the intrusion acts by handling the information from the activities, network logs, auditing data and something acquired through other ways. According to the differences of intrusion detection theory, it includes misuse-based detection and anomaly-based intrusion. Misuse detection is currently the most popular actual systems in the world[5].

However, the use of the existing intrusion detection algorithm based on model checking. The paper tries to solve the problems: Relative to detect intrusion attack, these algorithms ability, is not known at present, because of the lack of a kind of available data sets are available for this kind of algorithm implementation of detection.

It is the first step to choose the data sets to study and evaluate the various intrusion detection algorithm[6]. In the intrusion detection algorithm based on the pattern matching, the quality of the data set has direct influence on the functions of the detection system.Therefore,we need to have a recognized,excellent performance evaluation data set,otherwise,all kinds of algorithm and improved, there is no more basis and platform.

Currently recognized in academia is based on intrusion detection dataset MITLL collection, IDS by Columbia University laboratory organizing form of security audit dataset KDD CUP99 [7]. Many papers and research results are based on the basis of the data sets, but for the intrusion detection system based on formal methods no longer apply. In this paper, the structure of KDD CUP99 intrusion detection dataset, attack the distribution and feature selection, etc, on the basis of analysis, is constructed based on the behavior of intrusion detection data set.

Structure Data Set bases on the Behavior of Intrusion Detection

KDD99 data set is a network connection and system audit data.It is about data fields. To describe the data fields with 41 characteristics, the basic attributes of each feature contains which links, such as the continuous time, protocol type, bytes, etc.This paper focuses on the intrusion detection algorithm is based on the behavior.It's log requirement is an attack action,rather than for a network connection. Therefore ,we need to structure log , this log is base on attack action. With reference to the KDD99 data set, construct new log collection based on behavior, in order to evaluate the intrusion detection algorithm.

For model-checking intrusion based detection[8][9][10]，attackpatterns are described with LTL formulae and audit log is described with automata . In this way，model checking algorithms can be directly used to check whether event sets of records are inline with the attack patterns. Since

current model checking tools can effectively check the system states as high as 10^{120} . Compared with traditional pattern matching algorithms[11], the model checking method for large-scale network misuse detection is particularly effective . Figure 1 is the theory about Intrusion detection method based on model checking[12]:

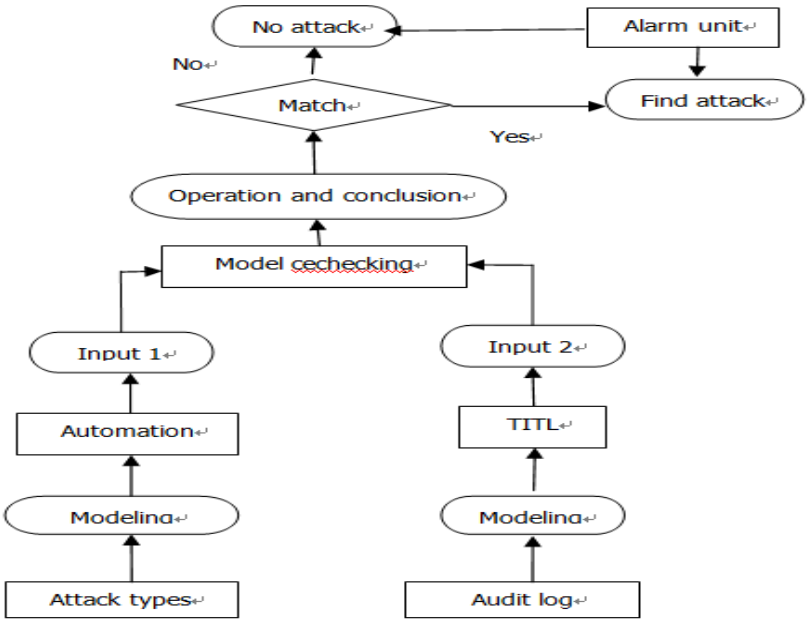

Fig.1.Intrusion detection method based on model checking

As show Fig.1 ,we can see that attack types in the first place in formal way described as logical formula as an input, and then the system audit records (that is the paper constructed based on the behavior of the data set) modeling as another input, and then run, finally came to the conclusion that whether or not found.

We use examples smurf to show the structure of the data set:

First step：Principle analysis Smurf attack, when being attacked host did not send the destination address for a subnet of the broadcast address of the packet, but the attacked host received a packet from all hosts in this subnet.

Second step：Smurf attack behavior decomposition，（1）To a broadcast address subnet, with a specific request, such as ICMP echo request package. (2) the source address camouflage to attacked the host address.

Third step：The attack is decomposed into action sequences, and build atomic formula for atomic motion model AP= { attacked.send, attacked.receive.i }. attacked.send is the destination address of the attacker to send packets broadcast address for a subnet. attacked.receive.i is Received the source address for a subnet of a host of packets.

Fourth step: According to the temporal relationship between the atomic motion, the atomic formula, can get describe smurf attack logic formula[13][14][15]:

$$G[\ (\neg attacked.send) \Rightarrow XF(\forall pi\ \ (attacked.receive.i))]^*$$

Fifth step: From logic formula describes attack behavior characteristics，we can get behavior field in the meaning of the data set,as show table 1.

op	source	des	type	fragmentoffset	totallength	port	time
receive	1	this	icmp	0	n	0	2015.7.18.08:20:152
receive	2	this	icmp	0	n	0	2015.7.18.08:20:160
receive	3	this	icmp	0	n	0	2015.7.18.08:20:167
receive	4	this	icmp	0	n	0	2015.7.18.08:20:399
receive	5	this	icmp	0	n	0	2015.7.18.08:20:458

Table.1.Smurf attack data set

Op：the main actions（The host receives the ICMP reply message）

Source：Refers to send a message of the source host identification number.

Des：To send a message destination host identification number.

Type：protocol type。

Fragmentoffset：data packet slice offset.

Totallength：The size of the data transmission.

Prot：The port number about attack.

Time：action time.

The performance evaluation of commonly used model checking algorithm

Now, we have LTL - MC, ITL - MC and RASL - MC three kinds of intrusion detection algorithm evaluation. The four types of attacks a certain capacity of data sets timeliness evaluation.

The Probe attacks are IP_sweep, Post and Mscan three types of attacks，Efficiency analysis as show Figure 2:

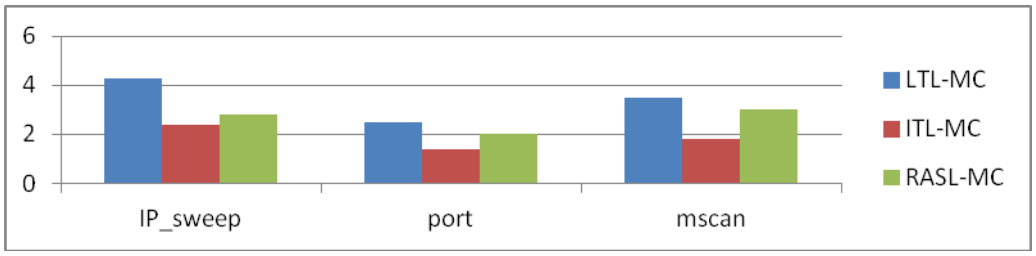

Fig.2. The probe of efficiency test results

DOS attack [15] include: including land, Neptune, pod, smurf, teardrop, apache, mailbomb, udpstorm8 kind of attack. Efficiency analysis as show Figure 3:

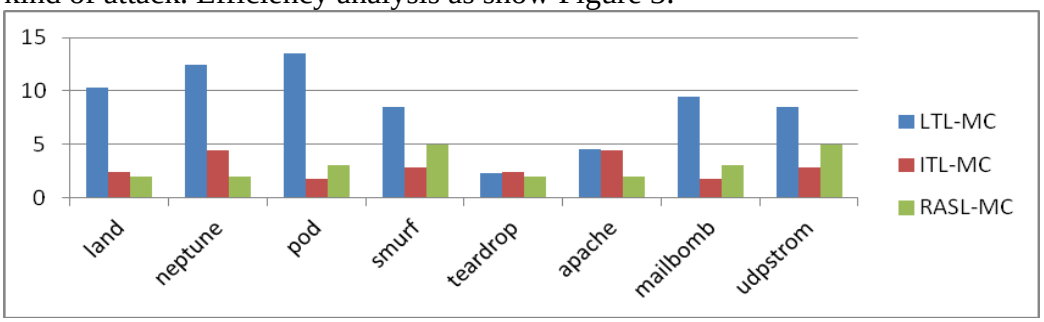

Fig.3. The DOS of efficiency test results

R2L attacks including FTP - write, PHF, imap, warezclient, warezmaster, xsnoop, sendmail. Efficiency analysis as show Figure 4:

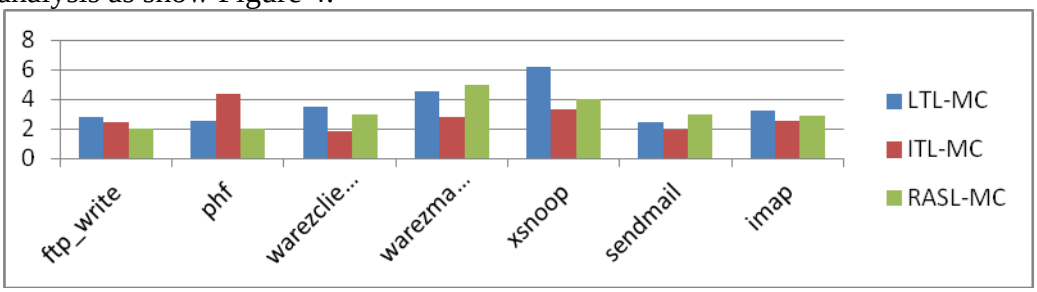

Fig.4. The R2L of efficiency test results

U2R attacks including rootkit, buf_overflow httptunnel, xterm. Efficiency analysis as show Figure 5:

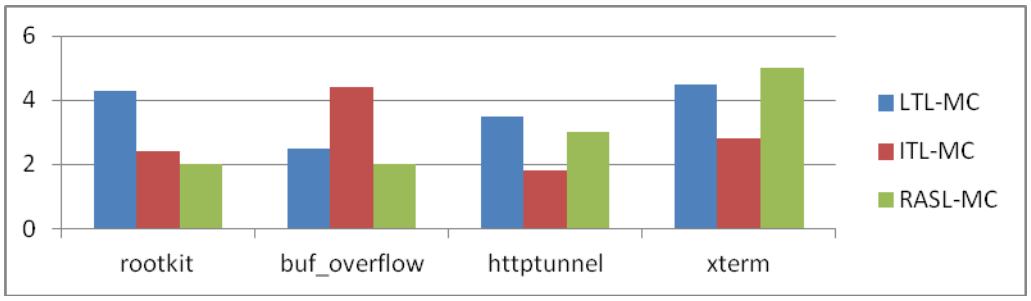

Fig.5. The U2R of efficiency test results

We can be seen from the diagram, ITL - MC efficiency is best, followed by RASL – MC[16], efficiency is the worst LTL - MC. Analysis the reason, is because of LTL - MC in describing nature ability not equal to is better than other two algorithms, therefore in the process of detection, relative efficiency is low.

Conclusion

Based on the analysis of the mechanical theory as the foundation, designed the soccer robot pick the ball institutions optimal design process, found aim function, select design variables and the corresponding optimization algorithm to optimize a complete set of institutions. At last through the test to get the final performance parameters of the institution. Experiments show that the system has higher accuracy and stability, the new optimize pick the ball have design basic requirements, and achieved good ideal control effect.

Acknowledgement

In this paper, the research was sponsored by the Nature Science Foundation of Henan Province (Project No. 201112400450401) and Youth Fund Project of Luoyang Institute of Science and Technology (Project No. 2010QZ16).

References

[1] Ali M Q, Al-Shaer E. Probabilistic model checking for AMI intrusion detection[C]//Smart Grid Communications (SmartGridComm), 2013 IEEE International Conference on. IEEE, 2013: 468-473.

[2] Schmerl S, Vogel M, König H. Using model checking to identify errors in intrusion detection signatures[J]. International Journal on Software Tools for Technology Transfer, 2011, 13(1): 89-106.

[3] J Olivain, J Goubault-Larrecq, The Orchids Intrusion Detection Tool, Proceedings of the 17th International Conference on Computer Aided Verification, Lecture Notes in Computer Science, 3576:286-290, Springer, Edinburgh, Scotland, UK, 2005.

[4] Guzzo A, Pugliese A, Rullo A, et al. Intrusion Detection with Hypergraph-Based Attack Models[M]//Graph Structures for Knowledge Representation and Reasoning. Springer International Publishing, 2014: 58-73.

[5] PATRICIA B. Model-checking timed temporal logics[J].Electronic Notes in Theoretical Computer Science,2009(231): 323-341.

[6] Xin you zhang hua shen zeng, Jia Lei. Research on intrusion detection dataset KDD CUP99. Computer engineering and design, 2010, 22:4809-4812 + 4816.

[7] Siddiqui, Mohammad Khubeb;Naahid, Shams.Analysis of KDD CUP 99 Dataset using Clustering based Data Mining.[J].International Journal of Database Theory & Application,2013,No.5.

[8] Yixian Yang, xin-xin niu. Intrusion detection theory and technology [M]. Higher education press, 2006.

[9] wei-jun zhu, qing-lei zhou, qin-xian zhang. Linear temporal logic model checking method based on DNA computing. Journal of computers, 2015.

[10] wei-jun zhu, qing-lei zhou. Time interval temporal logic and judgmental [J]. Journal of computer science, 2010, p.22.

[11] chancy. Network attack modeling based on temporal logic research [D]. Zhengzhou university, 2014.

[12] Zhao Yanke. Intrusion detection method based on temporal logic model validation study [D]. Zhengzhou university. 2014.

[13] W Zhu, Z Wang, H Zhang, A novel algorithm for Intrusion Detection based on Model Checking Interval Temporal Logci, China Communications, 8(3):66-72, 2011.

[14] Yang Lan.Design and Implemntation of Intrusion Detection System Based on Data Mining[J]. Energy Procedia, 2011, 13.

[15] Weijun Zhu; Qinglei Zhou; Weidong Yang; Haibin Zhang. A Novel Algorithm for Intrusion Detection Based on RASL Model Checking[J]. Mathematical Problems in Engineering, 2013.

[16] Wei-jun zhu. Time interval temporal logic model checking: theory, algorithms and applications [D]. Xi 'an university of electronic science and technology, 2011