ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

Multi-Agent Intrusion Detection System Based on Immune Principle

JingXu [1], SenXu[2], YongzhongLi[3]

Lecturer, School of Information Engineering, Yancheng Institute of Technology, Yancheng Jiangsu, China[1]

Associate Professor, School of Information Engineering, Yancheng Institute of Technology, Yancheng Jiangsu, China[2]

Professor, School of Computer Science and Engineering, Jiangsu University of Science and Technology, Zhenjiang Jiangsu, China[3]

ABSTRACT: Speed-bottleneck and bandwidth problems in traditional distributed intrusion detection system are solved by multi-agent, but false negative rate and false positive rate are still high. In order to solve the above problems, immune principle and multi-agent are combined, then immune agent is constructed and intrusion detection model based on immune agent is builded. An improved dynamic clonal selection algorithm is proposed. Operating principle of the system and realization of each agent are illustrated in detail. The proposed model and algorithm are simulated by KDD'99 datasets. Key parameters are optimized. Compared to other results, the proposed method has low false positive rate and higher detection rate in Dos, Probing, U2R, and R2L attacks.

KEYWORDS:Intrusion detection; multi-agent; immune principle; dynamic clonal selection algorithm; immune agent

I. INTRODUCTION

The internet has changed the traditional production, business, and lifestyle. It has brought us convenience, but network security issue is in a huge challenge. Intrusion detection system (IDS) plays an import part in network security defense, which is a proactive safety defense method [1]-[4].The IDS experienced four stages: host-based IDS, multi-host based IDS, network-based IDS, and distributed IDS. However, the traditional distributed IDS has network bandwidth problem [5].

An autonomous agent for intrusion detection (AAFID) was proposed [6]. In AAFID, nodes of the IDS are arranged in a hierarchical structure in a tree. Therefore, the network bandwidth problem can be solved. However, agents in AAFID are not mobile. When all collected data move to the central agent, it causes bottleneck of speed. In addition, they have bad real time capability and failure of single point because of the central processing node.

II. RELATED WORK

The aforementioned problems can be solved by utilizing the mobile characteristics of agents and its distributed collaborative calculation capability [7]-[9]. In [10], an IDS model based on mobile agentwas proposed. The system consists of data collection agent, data analysis agent, and system fixed agent. Data collection agents obtain information about each node by roaming. Data analysis agents exchange information with different data collection agents. System fixed agents are responsible for taking necessary measures to invasion. Extensibility of the system is improved and single-point failure of the system is solved by the mobility of the mobile agent.

An automatically optimized distributed IDS based on mobile agent was proposed [11]. It is composed of sensor agent, analyzer agent, and console and improver agents. The sensor agent can send data packets intelligently to the analyzeragent, whichis most suitable to handle the batch data. Analyzer agent is optimized to ensure the efficiency of the system by utilizing the mobile characteristic oftheconsole and improver agent.

10.15680/ijircce.2015.0304001

Although [10] and [11] have solved failure of single point and bottleneck problems and improved design and structure of the IDS, the false positive rate and false negative rate are still high.Biologic immune system realizesidentification and classification by learning, memory, and extraction, which is similar to the IDS. Therefore, the biological immune principlecan be applied into the IDSto improve detection performance [12].

Inspired by the immune theory, Kim and Bentley proposed a dynamic clonal selection algorithm for theIDS [13]. It can increasetrue positive (TP) rate, reduce false positive (FP) rate. However,memory detectors can not tolerate integrated self-sets in the algorithm and have infinite long-life cycle. In addition, co-stimulatory is needed when mature detectors detectantigens.

Both mobile agent technology and immune principle are distributed systems, whichhave multiple independent entities and are capable of self-learning. Moreover,they can respond according to the changeof external environment. Therefore, immune principle and mobile agent technology can be combined. The model of mobile agent IDS based on immune principle is proposed. The model is not only scalable and adaptive, but also improves the detection performance. The immune agents are constructed. These immune agents are dynamic, so they can patrol in the system to detect intrusion. It is similar to the immune cells in biology.

III. PROPOSED MODEL

The proposed system is consisted of different agents, and all agents in the system can be divided into four categories according to their functions: control agent, acquisition agent, detecting agent, and response agent. The model is shown in Fig. 1, where the abbreviation MAE refers tomobile agent environment. Agents collaborate with each other to simulate the function of the immune system to detect intrusion.

Control agent: it is primarily responsible for the management, coordination, and control agent on host, which is monitored. It is similar to bone marrow and thymus in biological immune system. It creates and dispatches detecting agents after receiving the signal from acquisition agents. After detecting agents are created, their work isindependent of control agent. Even if the control server is attacked, it will not affect the work of the detecting agents, which have been generated. Detecting agents can clone and move to the destination host in order to detect intrusion. Therefore, the failure of single point problem can be eliminated.

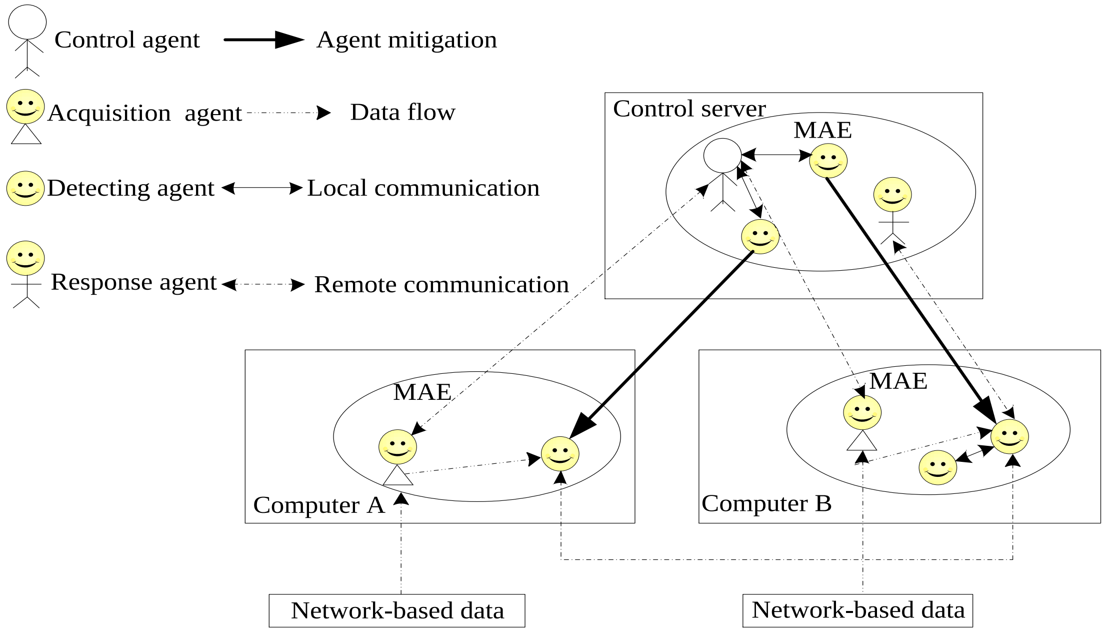

Fig.1 Architecture of the system

Acquisition agent: data acquisition is the base of intrusion detection. Detection, analysis, and response are built on the basis of data acquisition. The data sources ofIDS include host-based data and network-based data. Acquisition agents are distributed on major nodes in network, which are responsible for intercepting network data packets and pretreating the intercepted network data packets. As the gathered data is very large, the related information should be filtered by acquisition agents in order to delete the useless information.

Detecting agent: The data sent by acquisition agent is detected and analyzed by the detecting agent. Detectors are composed of immature detectors, mature detectors, and memory detectors. The immature detectors are generated initially, which are not processed by negative selection. The mature detectors are survived after negative selection. The memory detectors simulate secondary response mechanism in immune system, which detect the appeared attacks. The generated mature and memory detectors are embedded in the agents, being detecting agents. Detecting agents are divided into B-agent and M-agent. Each B-agent has a detector set, where all detectors are mature detectors by training. Each M-agent is also has a detector set, where all detectors are memory detectors by training. Fig. 2 shows the generating process of the detecting agent. The detecting agents can mitigate among different computers and complete intrusion detection by communicating and collaborating with each other.

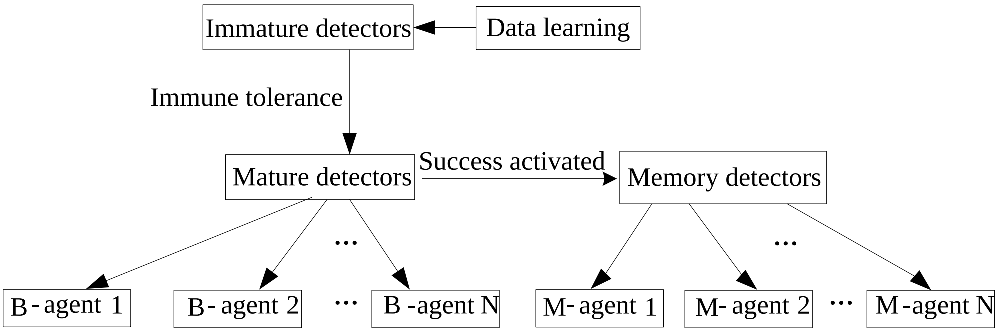

Fig. 2 Generating process of the detecting agent.

Response agent: If B-agent or M-agent has detected or suspicious intrusions, the response agents will be activated to send alarm and respond.

IV. PROPOSED ALGORITHM

A.The proposed dynamic clone selection algorithm

The detailed steps of the improved dynamic clone selection algorithm are as follows.

Step 1: Immature detectors are generated randomly. They are compared with the given self set by the negative selection algorithm, and then the matched detectors are deleted. Finally, the new immature detectors are compensated till the maximum size of a non-memory detector population. The same process continues in differentgeneration of the tolerization period T. When the total generation reaches T, the immature detectors, whose generation is equal to T, become mature detectors.

Step 2: In the generationof T+1, mature detectors detect new antigen set. If one antigen matches with the mature detector, the antigen will be determined whether it belongs to the current self set. If yes, the mature detector will be deleted, and the antigen will be added into the current self set. Otherwise, the antigen will be deleted, and the counter of the corresponding mature detector pluses one. If the value of the counter exceeds the activation threshold A, the mature detector will enter the memory detector set. If it is less than A and the age of the mature detector is larger than the life cycle L, the mature detector will be deleted. Or else, the detection process will go on.

Step 3: In the generation of T+2, when the memory detector matches with an antigen, the antigen will be determined whether it belongs to the current self set. If yes, the antigen will be deleted. If not, the memory detector will be deleted. The remaining antigens are compared to the mature detector population, and the detection process is the same as that in the generation of T+1.

Step 4: From the generation of T+3, the detection process is the same as that in the generation of T+2. The process will be continued to detect the changing antigen population till the system is closed.

B. Abnormal detection

Detecting work is completed by acquisition agent, B-agent, and M-agent. At the beginning, M-agent is null. The acquisition agent obtains log information on the host and data packets transmitted on the protected network. The obtained data is pretreated by the acquisition agent, and the binary antigen is created.

The detector and intrusion detection are generated by incomplete matching rules. Only if the protocol and service-type of the data to be detected match with the detector, the matching process will start. When the number of matching characteristic is no less than r, the data is considered as an intrusion. Otherwise, it is normal.

In order to improve the detecting efficiency of the appeared intrusions, the detector in B-agent is upgraded to the memory detector when the number of non-self antigens in the detecting period exceeds the threshold A. After that, the memory detector is dispatched to the M-agent. The detailed abnormal detection flow chart is shown in Fig. 3.

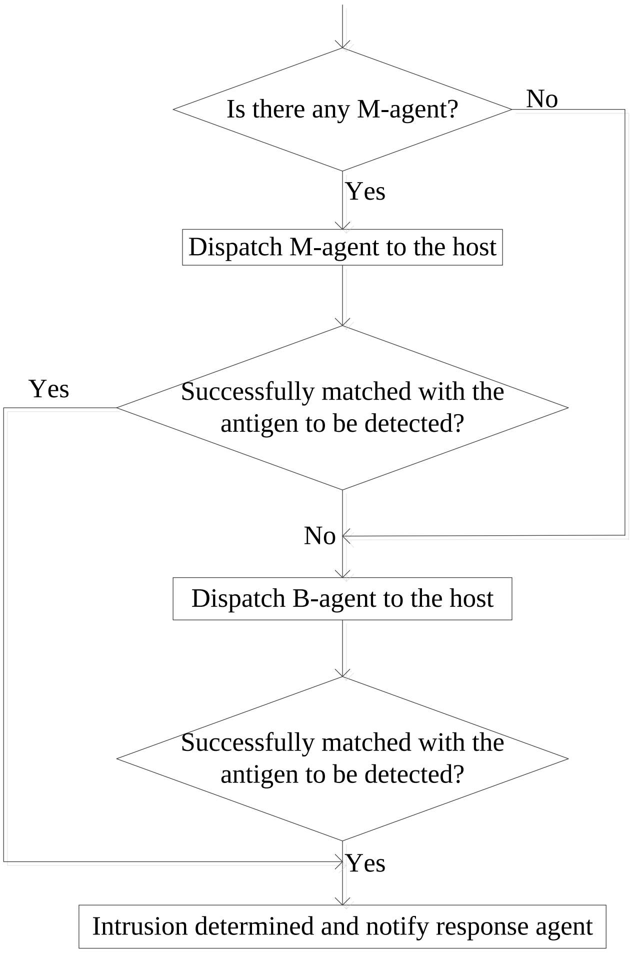
Is there any M-agent?
No
Yes
Dispatch M-agent to the host
Yes
Successfully matched with the antigen to be detected?
No
Dispatch B-agent to the host
Successfully matched with the antigen to be detected?
Yes
Intrusion determined and notify response agent

Fig. 3Detailed abnormal detection flow chart.

C. Abnormal response

After receiving the abnormal signal, the response agent moves to the suspected host and proceed with the abnormal condition. During the process, human-computer interaction is also focused on. If the administrator confirms the intrusion, interruption will be taken immediately, such as discarding a suspicious data package, cutting a suspicious connection, locking the account, restricting login, and so on. If administrators discover this intrusion is unknown, then the corresponding detector is added into the memory detector set and dispatched to the M-agent. In addition, if the administrator confirms false alarm, the corresponding detector will be deleted from the B-agent or M-agent, which is added to the self-antigen set.

V. Simulation Results

KDD Cup'99 dataset is adopted as the training data. Taking Dos-attacks as an example, the dataset is constructed as follows.

(1) The back, land, neptune, pod, smurf, and teardrop of Dos-attack records are 1000, which are randomly and proportionally obtained from kddcup.data_10_percent.

(2) The normal data is 1000, which is randomly selected from kddcup.data_10_percent.

(3) The back, land, neptune, pod, smurf, teardrop, processtable, and apache2 of the Dos-attack records are 1000, which are proportionally selected from the corrected file, and the normal data from the corrected file is 1000.

The selected Dos-attack records and normal data are combined, which are divided into training data and detecting data according to the proportional rate of eighty and twenty percent, respectively.

A.Parameter r influenceon the performance of the system

In the learning stage, when the value of r is smaller, the number of mature detectors in system is less. In addition, it results in low TP and FP rates.Likewise, when the value of r is larger, more mature detectors, which contain self-character, will be produced. Moreover, it results in high FP rate.

Therefore, not only effective formation of mature detectors but also FP rate should be considered when the r parameter is set. When A equals 5, T equals 5, L equals 10, N equals 1, and the value of iteration is 100, TP and FP rates changing with r is given in Fig. 4. The simulation results are consistent with the above analysis. Therefore, r is selected as 6.

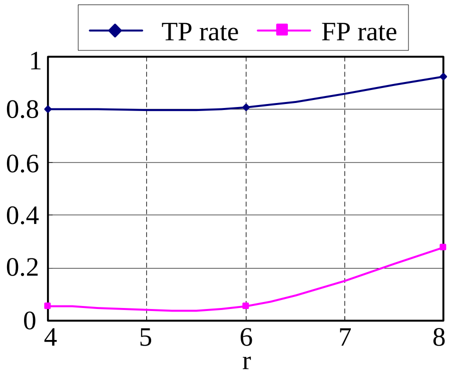

Fig. 4 Chart of TP and FP rates when r varies.

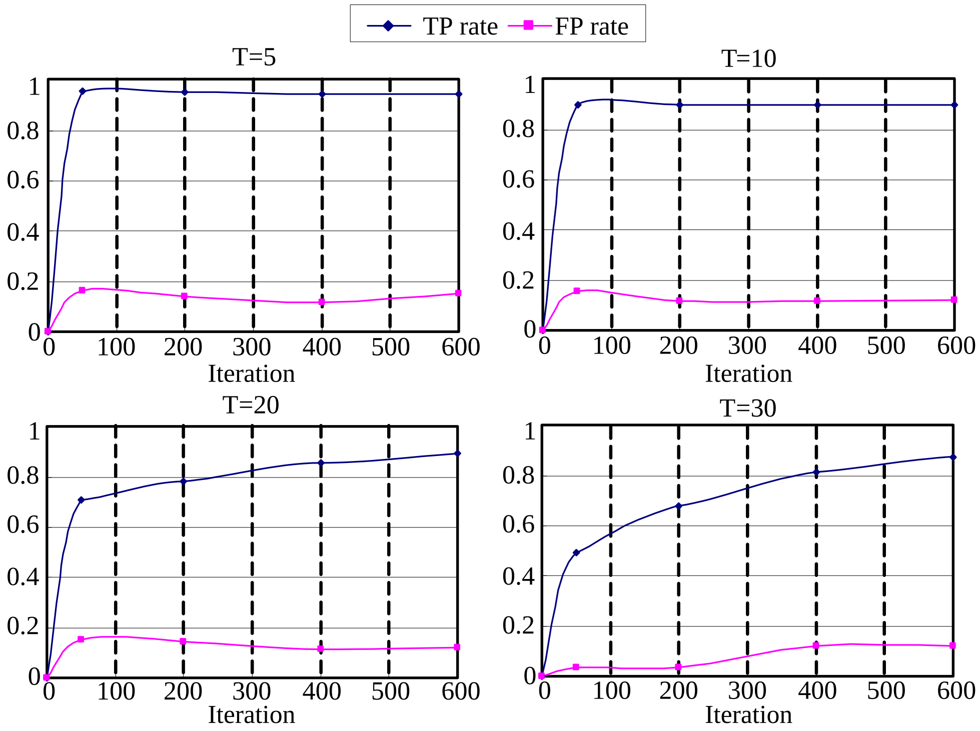

Fig.5 TP and FP rates when T varies.

B.Parameter T influenceon the performance of the system

When thetolerance period of the self-antigen reaches T,immature detector becomes mature detector. The number of immature and mature detectors is constant value.Thus, when T is larger, the number of mature detectors is less. TP and FP rates will be reduced. When T is smaller, more mature detectors will be produced, and it results in high TP rate. Because of smallerT, some of the mature detectors will not tolerate completely and FP rate will be higher.

WhenA equals 5, L equals 10, and r equals 6,TP and FP rates changing with T is given in Fig. 5. From Fig. 5, T increase, TP and FP ratesfall. Therefore, T is chosen as 10.

C.Parameter L influenceon the performance of the system

Parameter L represents the lifecycle of mature detector. The value of L directly affects the number of mature detectors, thereby affecting the TP and FP rates of the system. When L is bigger, more mature detectors will be created, and TP and FP rates will increase. Otherwise, TP and FP rates will fall.

WhenA equals 5, T equals 10, and r equals 6, TP and FP rates changing with Lis shown in Fig. 6.From Fig.6, with the increase of L, TP and FP ratesrise, but TP rate increases slightly.After L reaches 20, TP and FP rates isbasically stabilized. Therefore, parameter L is selected as 20.

10.15680/ijircce.2015.0304001

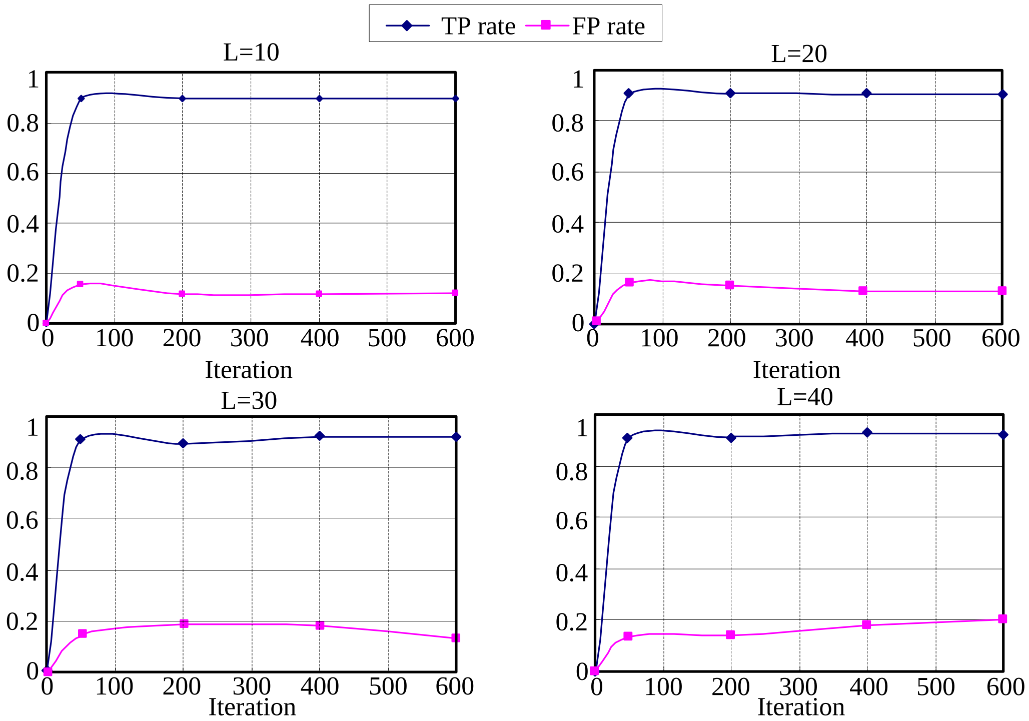

Fig. 6 TP and FP rates when L varies.

D.Parameter Ainfluenceon the performance of the system

Parameter A represents the activation threshold of mature detector. Since A must be smaller than L, A is taken 5 and 15 in both cases.When T equals10, Lequals 20, and r equals 6,TP and FP rates changing with Ais given in Fig. 7.

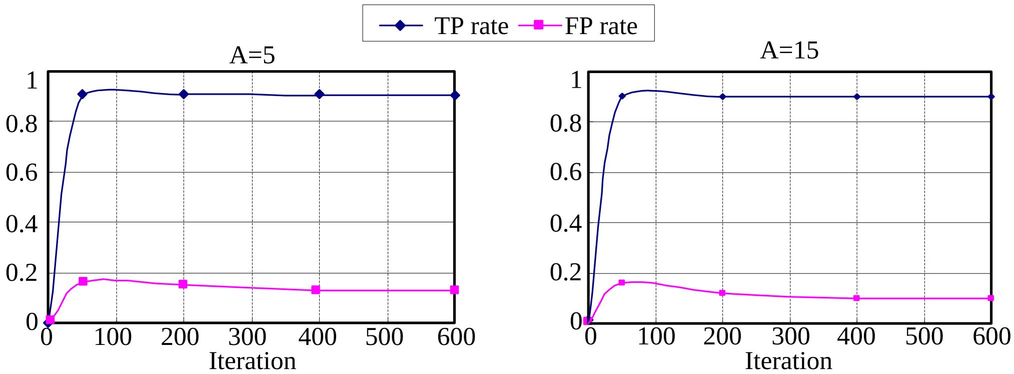

Fig. 7 TP and FP rates when A varies.

As can be seen from Fig. 7, when A increases, FP and TP rates decrease. Increasing A, the frequency of mature detector activation will be reduced. Therefore, in the fixed iteration, less memory detectors will be generated. FP rate, which caused by memory detectors, will be reduced. Considering L, A is taken as 15.

E.Comparison among other methods and the proposed method

In simulation, the training data set is selected from the data set of kddcup.data_10_percent, while the test data set is obtained from the corrected data set, as shown in Table 1. The number of non-memory detector is 100000. In addition, the training data set is divided into four self-antigen and four nonself-antigen sets. The maximum iteration is 200. The simulation is repeated five times, and the results are averaged.

Table 2 gives the simulation results compared with different methods. From Table 2, the proposed method has good TP rate in Normal, Dos, and Probing attacks. The TP rate in U2R attacks is significantly improved compared with other methods. However, the TP rate in R2L attacks in the proposed method is less than that in Ref. [13]. The main reasons of the above results are as follows: in training data set, the number of samples of the U2R and R2L attacks is far less than that of Normal, Dos, and Probing attacks, and then stable knowledge cannot be generated. Therefore, most of the U2R and R2Lattacks are judged as normal data. In test data set, the number of samples of the R2L attacks is relatively large, which results in the small TP rate.

Table1 Number of samples in training and test sets

Type of collection	Number of samples	
	Training set	Test set
Normal	97278	60593
Dos	391458	229853
Probing	4107	4166
U2R	52	228
R2L	1126	16189

Table2 Comparison among other methods and the proposed method

Type of collection	TP rate		
	Ref. [13]	Ref. [5]	The proposed method
Normal(60593)	99.5%	96.2%	96.53%
Dos(229853)	97.1%	96.6%	97.82%
Probing(4166)	83.3%	80.9%	90.71%
U2R(228)	13.2%	9.6%	73.68%
R2L(16189)	8.4%	0.1%	4.17%

VI. CONCLUSION AND FUTURE WORK

Combined with mobility and collaborative of the agents and the immune theory,the IDS model based on immune agent has been proposed.An improved dynamic clonal selection algorithm is illustrated.The proposed algorithm and model are simulated byKDD'99 datasets.Simulation results show thatthe proposed method has low FP rate and improves the TP rate in Dos and Probing attacks compared with other methods. Moreover, the TPrate of U2R attacks is greatly improved. For R2L attacks, the TP rate compared with Ref. [5] is increased, but still low.

Future work is to improve the TP rate in all kinds of attacks, especially R2L attacks. In addition, the FP rate should be reduced. The types of attacks should be detected in detail.

REFERENCES

1. J. McHugh, A. Christie, and J. Allen, ‘Defending Yourself: The Role of Intrusion Detection Systems’, IEEE Software, Vol. 17, Issue 5, pp. 42-51, 2000.
2. G.Creech and J. Hu,‘A Semantic Approach to Host-Based Intrusion Detection Systems Using ContiguousandDiscontiguous System Call Patterns’, IEEE Transactions on Computers, Vol. 63, Issue 4, pp. 807-819, 2014.
3. C.-J. Chung, P.Khatkar, T. Xing, J. Lee,and D. Huang,‘NICE: Network Intrusion Detection and Countermeasure Selection in Virtual Network Systems’, IEEE Transactions on Dependable and Secure Computing, Vol. 10, Issue 4, pp. 198-211, 2013.
4. Y. Zhang, L. Wang, W. Sun, R.C.Green, and M.Alam,‘Distributed Intrusion Detection System in a Multi-Layer Network Architecture of Smart Grids’, IEEE Transactions on Smart Grid, Vol. 2, Issue 4, pp. 796-808, 2011.
5. Y. Liu, D. Tian, X. Yu, and J. Wang,‘Large-Scale Network Intrusion Detection Algorithm Based on Distributed Learning’, Journal of Software, Vol. 19, Issue 4. pp. 993-1003, 2008.

6. E. H. Spafford and D. Zamboni, 'Intrusion Detection Using Autonomous Agents', Computer Networks, Vol. 34, pp. 547-570, 2000.
7. S.Fenet and S.Hassas,'A Distributed Intrusion Detection and Response System Based on Mobile Autonomous Agents Using Social Insects Communication Paradigm',First International Workshop on Security of Mobile Multiagent Systems, Montreal: Electronic Notes in Theoretical Computer Science,Vol. 63, PP. 43-60, 2001.
8. A.Neda andA.Reza,'MAIS-IDS: A Distributed Intrusion Detection System Using Multi-Agent AIS Approach', Engineering Applications of Artificial Intelligence, Vol. 35, pp. 286-298, 2014.
9. C.M.Ou, 'Multiagent-Based Computer Virus Detection Systems: Abstraction FromDendritic Cell Algorithm With Danger Theory', Telecommunication Systems, Vol. 52, pp. 681-691, 2013.
10. R. Wang, H. Wang, and X. Xu,'Research on Intrusion Detection System Based on MobileAgent',Journal on Communications, Vol. 25, Issue 1, pp. 22-29, 2004.
11. J. Wang, D. Li, and D. Feng,'An Automatically Optimized Distributed Intrusion Detection System Using Mobile Agent',Journal of Computer Research and Development, Vol. 43, Issue 1, pp. 9-14, 2006.
12. R. Pokrywka, 'AritificialImmune System Framework for Pattern Extraction in Distributed Environment', The International Joint Conference on Software Technologies,Vol. 1, pp.179-183, 2010.
13. J. Kim andP. Bentley,'Towards an Artificial Immune System for Network Intrusion Detection: An Investigation of Dynamic Clonal Selection', The Proceeding of the Congress on Evolutionary Computation, Vol. 2, PP. 1015-1020, 2002.

BIOGRAPHY

Jing Xuis a Lecturer in the School of Information Engineering, Yancheng Institute of Technology, Yancheng, China. She received Master of Computer Application (MCA) degree in 2009from Jiangsu University of Science and Technology, Zhenjiang, China. Her research interests are computer networks and network safety.

Sen Xuis anAssociate Professor in the School of Information Engineering, Yancheng Institute of Technology, Yancheng, China. He received Ph. D. of Computer Application degree in 2010from HarbinEngineeringUniversity, Harbin, China. Her research interests are computer networks and network safety.

Yongzhong Liis a Professor in the School of Computer Science and Engineering, Jiangsu University of Science and Technology, Zhenjiang, China. He received Master of Communication and Electronic Systems degree in 1995from University of Electronic Science and Technology, Chengdu, China. Her research interests are computer networks and network safety.