

# A Novel Distributed Intrusion Detection Model Based on Immune Mobile Agent

Yongzhong Li, Rushan Wang, Jing Xu

School of Computer Science & Engineering Jiangsu University of Science and Technology Zhenjiang, China  
liyongzhong61@163.com

**Abstract**—Intelligent and distributed is a development direction of intrusion detection system in future. However, current distributed intrusion detection system mostly use distributed component to collect data that are sent to processing center. Data is analyzed in the processing center. Nevertheless, these models have the following problems: bad real time capability, bottleneck, and single point of failure. In order to overcome these shortcomings, a new distributed intrusion detection model based on mobile agent is proposed in this paper. Intelligent and mobile characteristics of the agent are used to make computing move to data. Analysis shows that the network load can be reduced and the real time capability of the system can be improved with the new model. The system is robust and fault-tolerant. Because mobile agent only can improve the structure of system, dynamic clonal selection algorithm is adopted for reducing false positive rate. The simulation results on KDD99 data set prove that the new model has low false positive rate and high detection rate.

**Index Terms**—distributed intrusion detection, mobile agent, immune agent, network security

## I. INTRODUCTION

Traditional intrusion detection systems(IDS) are centralized and based on a monolithic architecture. Data are collected on a single machine by looking at log files or network flow and are analyzed on a single computer, which has some defects both in the structure and in the detection technology. So distributed intrusion detection system (DIDS) appears. It becomes a research focus in the field of intrusion detection. Reference [1] presented a distributed information-gathering step, but centralized on analyzing process. The Graph-based Intrusion Detection System (GrIDS) [2] and Event Monitoring Enabling Responses to Anomalous Live Disturbances (EMERALD) [3] are IDS that use a hierarchical approach in a more sophisticated way. The hierarchical approach seems to show better scalability by allowing local analyses at distributed monitoring areas. However, a monitor operating at the highest level may induce single point of failure. When the topology of current network is changed, it causes a change of network hierarchy, and the whole mechanism for aggregation of local analysis reports must be changed. Autonomous Agent for Intrusion Detection (AAFID) is the first attempt to use autonomous agents for network intrusion detection by Spafford and Crosbie in [4]. In AAFID, nodes of the IDS are arranged in a hierarchical

structure in a tree. Agents in AAFID were not mobile. Current DIDS mostly use distributed component to collect data, and then send collected data to processing center. These models solve the problem of distributed data acquisition effectively in wide bandwidth network. However, they have bad real time capability, bottleneck problem, and single point of failure because of the central processing node. The above-mentioned problems can be solved by utilizing the intelligent, mobile, and self-adaptive characteristics of agent and its distributed collaborative calculation capability [5] – [6].

False positive rate and false negative rate are other import aspects that IDS must consider. In [7], the authors stated similarities between the defenses of natural immune systems and computer security: both must discriminate self and non-self to protect a complex system from inimical agent. Be inspired of immune system, Kim and Bentley in [8] have proposed dynamic clonal selection algorithm and shown that this algorithm could reduce false positive rate.

In this paper, dynamic clonal selection algorithm is adopted. Detectors are embedded in agents. With their communication mechanism, detection agents can detect cooperatively. Using the mobile characteristic of agent, detection agent can move to local host, and thus it can reduce network load and improve real-time capability. The model is full distributed.

## II. RELATIVE KNOWLEDGE

### A. Immune system

The immune system [9]–[12] is a complex network of organs and cells responsible for the organisms defense against alien particles. Among a large number of different innate and acquired cells, lymphocytes play a central role. Lymphocytes are classified into two main types: B-cells and T-cells. B-cells are antibody-secreting cells and T-cells kill antigens or help or suppress the development of B-cells. Both originate from bone marrow, and they are developed by the bone marrow and the thymus, respectively.

Before leaving the bone marrow and the thymus, maturing B- and T-cells have to pass the last test-negative selection. Mature B- and T-cells that pass the negative selection are released from the bone marrow and thymus, respectively. The development of B-cells and T-cells are shown in Fig. 1.

The antibodies of B-cells, which recognize harmful antigens by binding to them, are activated directly or indirectly. When B-cell antibody receptors bind to antigen epitopes with strong affinity above a threshold, they are

---

This paper is supported by Research fund University of Jiangsu Province and Jiangsu University of Science and Technology's Basic Research Development Program (No. 2005DX006J)

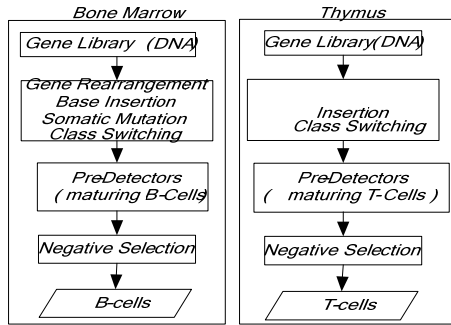


Figure 1. Development of B-cells and T-cells

directly activated. When B-cell antibody receptors fasten to antigen epitopes with weak affinity, MHC molecules try to find some hidden antigens inside cells. When MHC molecules find them, they transport them on the surface of B-cells. The receptors of T-cells are genetically structured to recognize the MHC molecule on the B-cell surface. When the T-cell binds to an MHC molecule with strong affinity, it sends a chemical signal to the B-cell, which allows it to activate, grow, and differentiate.

Clonal selection is immediately followed with B-cells activate. The more specific antigens B-cells bind to, the more chance being selected for cloning they have. When antigens activated B-cells, they produce memory cells for the reoccurrence of the same antigens in the future. Therefore, the secondary response is quickly.

### B. Mobile agent

Mobile agent is a type of software agent, with the feature of autonomy, social ability, learning, and most import, mobility [13].

Mobile agent can transport its state from one environment to another with its data intact and still be able to perform appropriately in the new environment. When a mobile agent decides to move, it saves its own state, transports this saved state to the next host and resumes execution from the saved state. Fig. 2 illustrates the work processing of mobile agent.

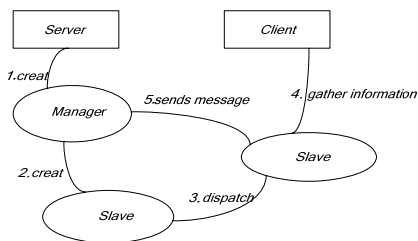


Figure 2. Work processing of mobile agent.

Mobile agent has many advantages. Mobile agent makes computation move to data, it can reduce network load. Because the actions are dependent on the state of the host environment, it is dynamic adaptation. It can operate without an active connection between client and server, so it has the capability of faults tolerance to network.

Mobile agent neither brings new method to detect for IDS nor increases detection speed for some kind of attracting. Nevertheless, it improves the design, construct, and execute of IDS obviously.

## III. THE INTRUSION DETECTION MODEL BASED ON IMMUNE MOBILE AGENT

### A System architecture

Be inspired of immune system, this paper combines immune mechanism with mobile agent, and constructs some immune agents to monitor and detect attraction on the network. Each immune agent can be regarded as immune cell. Like B-cells and T-cells circulating around the body in the blood and preventing the body by suppressing or killing the foreign invaders, immune agents roam on the network, and monitor and detect attacking.

Fig.3 presents the architecture of intrusion detection based on immune mobile agent. It composes of central control agent (C-agent), detection agent (B-agent), memory agent (M-agent), and response agent (K-agent). C-agent runs in the server and plays a role of manger. B-agent and M-agent travel though the network in order to detect attacking. If any attacking is detected by B-agent or M-agent, K-agent is activated and responds to it immediately.

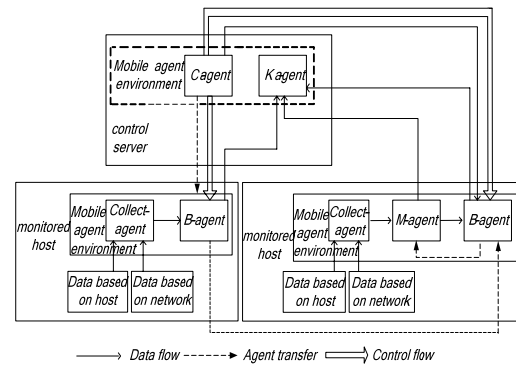


Figure 3. Architecture of immune agent intrusion detection system.

C-agent is a kind of agents, which mainly manage, coordinate, and control roaming agent on the network. Its function is similar to that of bone marrow and thymus. It can create, dispatch, and recall agent. Once B-agent is created, it can work continually without the connection between server and client. Although we adopt server and client model, it does not induce single point of failure.

Each B-agent contains a set of mature detectors. The function of B-agent is similar to that of B-cells. B-agent strays on the network to monitor intrusion. If antigen comes, B-agent is activated, and it will move to the local host to detect whether intrusion occurs.

Each M-agent contains a set of memory detectors. It imitates the mechanism of secondary response in immune system. If antigen comes and M-agent exists, M-agent is activated, and they will be detected by M-agent firstly. If it does not match these antigens, B-agent will detect continually. It can improve the speed of detecting known intrusion.

The function of K-agent is analogous to that of T-cells. If any intrusion is detected by B-agent or M-agent, K-agent will be activated immediately. It will act in response to it by disconnecting suspicious connection, locking account, or restricting login.

Collect-agent's main job is collecting data, which are the foundation of intrusion detection system. It can collect data based on host and based on network. Collect-agent in this paper mainly captures network packet. In order to improve efficiency of detection, collect-agent needs to extract useful property of data packet besides of capturing data packet.

### B. Generation of detectors

Detectors play an important part in intrusion detection. The more attacking features these detectors have, the higher detection rate the system has. The less normal network features these detectors contain, the less false positive rate the system has. Kim presents dynamic clonal selection algorithm and experiment shows it can reduce false positive rate with high detection rate [8]. When detectors are generated, they are embedded in mobile agent. Suppose that there are  $N$  mature detectors in total and each B-agent can carry  $n$  detectors, the system will generate  $\lceil N/n \rceil$  B-agents. These agents with detectors roam on the network and realize the distributed computing.

The dynamic selection algorithm is described in detail as follows:

**Step1:** create an immature detector population with random detectors, and perform negative selection by comparing immature detectors to self-antigens. Because of negative selection, the immature detectors binding any self-antigen are deleted from the immature detector population and then new immature detectors are generated until the number of immature detectors becomes the maximum size of a non-memory detector population. The same processes continue for the tolerance period ( $T$ ) number of generations. When the total number of generations reaches  $T$ , some immature detectors whose age reaches  $T$ , which were born at generation 1, become mature detectors.

**Step2:** at generation  $T+1$ , mature detectors start to detect new antigen set. If any mature detector matches an antigen, the match count of a mature detector increases by one. After all the given antigen are compared with all the existing mature detectors, the system will check whether the match counts of mature detectors are larger than a pre-defined activation threshold ( $A$ ) and whether the ages of mature detectors meet a pre-defined life span ( $L$ ). If the match count of a mature detector is larger than  $A$ , it becomes a memory detector. If the age of a mature detector meets  $L$ , the mature detectors are deleted from the mature detector population. Any antigens detected by activated mature detectors are deleted, and the remaining antigens are presented to the immature detectors for negative selection. If the number of existing immature detectors and mature detectors is less than the maximum size of a non-memory detector population, generate immature detectors with random detectors until they are equal.

**Step 3:** at generation  $T+2$ , when memory detectors match any antigen and the detected antigen bind any self-antigen, the memory detector are added to immature detectors. In addition, if the detected antigen does not

bind self-detectors, it is removed directly. The remaining antigens are matched by mature detectors and the process is the same as the period of  $T+1$ .

## IV. SIMULATIONS

In order to survey and evaluate research in intrusion detection, KDD99 data set is the data set, which was obtained from the 1998 DARPA. The data set is composed of a big number of connection records. Each connection is labeled as either normal or as an attack with exactly one specific attack type. Attacks in the data set can be classified into four main categories namely Denial of service (DOS), Remote to User (R2L), User to Root (U2R) and Probe. In our experiment, we only used 10 percent of the raw training data (kddcup.data\_10\_percent) for training and the test data set (corrected.gz) for testing. It is important to note that the test data is not from the same probability distribution as the training data, and it includes specific attack types not in the training data. The test data contains of 20 types of training attack and 17 types unknown attacks. The 37 types attacks are classified into four categories as follows:

**DOS:** {back, land, Neptune, pod, smurf, teardrop, processtable, udpstorm, mailbomb, apache2}

**R2L:** {ftp\_write, guess\_passwd, imap, multihop, phf, warezmaster, sendmail, xlock, snmpguess, named, xsnoop, snmpgetattack, worm}

**U2R:** {buffer\_overflow, loadmodule, perl, rootkit, xterm, ps, httptunnel, sqlattack}

**Probing:** {ipsweep, nmap, portsweep, satan, mscan, saint}

For each connection, there are 41 features. Among them, 32 features are continuous variables and 9 features are discrete variables. Among these features, some are redundant and some contribute little to the intrusion detection process [13]–[14]. Considering efficiency, we select features 1, 2, 3, 5, 6, 7, 8, 9, 15, 16, 18, 19 and 20 to compose of detector and choose statistical features 20 to 41 except for 30 to be collaborative signal. We can use much less, less, normal, more and much more to express bytes from source to destination. Moreover, the value of them is 000, 001, 010, 011, and 100, respectively. Fig. 4 shows its membership functions.

$$p(|x - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2} \quad (1)$$

According to Chebyshev's inequality (1) and the proportion of normal and attack data in KDD99, the

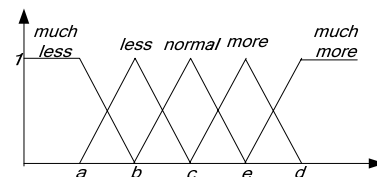


Figure 4. Features of detector.

values of variables in Fig. 4 are shown as follows:  $a = \mu - 2\sigma$ ,  $b = \mu - \sigma$ ,  $c = \mu$ ,  $e = \mu + \sigma$ ,  $d = \mu + 2\sigma$ . For all collaborative signals, we use normal, suspicious and

abnormal to express them, and the process is the same as the above. The value of it is 00,01 and 10 respectively.

This paper implemented on IBM's aglet toolkit, which is composed of a set of java-based mobile software agents that carry out specific tasks on the network and collaborate on the problem of network intrusion detection. Aglet is installed in three computers. Among them one is as sever, and others are as clients. When detectors are generated, they are embedded in agents. Utilizing aglet, agents can be dispatched, cloned, and recalled.

### B. Simulation results

- Test of Robust and Fault-tolerant of the system

When the system is start-up, B-agent in one host is broken off in order to observe its effect to the system. Experiment shows that system can discover the invalidated agent and then create and dispatch new agent to this host. One node invalidate does not induce disability of the system. This indicates that the system is robust and fault-tolerant.

- Detection result

The size of non-memory detectors is defined as 100000, the training data is divided into self-antigen set and non-self antigen set. In addition, in our experiment, self-antigen set and non-self antigen set are classified into four antigen clusters. Moreover, the iterative generation is set 200.

In table I, comparing with winning entry of KDD'99 Classifier Learning Contest, the proposed approach has a good performance in detecting DOS, Probe, U2R attack and Normal behavior. Nevertheless, the performance of detecting R2L is poor. This is because the packet of R2L is slightly different from the packet of normal. How to improve the ability of the detecting R2L and U2R is the future work.

TABLE I. COMPARISON WITH WINNING ENTRY OF KDD'99 CLASSIFIER LEARNING CONTEST

	Detection Result	
	TP of the winning entry	TP of the proposed approach
Normal(60593)	99.5%	98.127%
DOS(229853)	97.1%	97.565%
Probe(4166)9	83.3%	90.494%
U2R(228)	13.2%	71.491%
R2L(16189)	8.4%	0.371%

### V. CONCLUSIONS

In this paper, a new distributed intrusion detection model based on immune mobile agent is proposed. Dynamic clonal selection algorithm and mobile agent are described in detail. The simulation results showed that our model is efficiently to classify the anomaly profile from the normal profile. Our model has following advantages. First, the model realized that computing move to data by utilizing mobile agent. Therefore, real time capability is improved and bottleneck problem is overcome. Second, compared with other hierarchical model, it surmounts single point of failure. Dependability

of the system is enhanced. In addition, the system is robust and fault-tolerant. Third, false positive rate is low and true positive rate is high by adopting dynamic clonal selection algorithm.

### ACKNOWLEDGMENT

This paper is supported by Research fund of University of Jiangsu Province and Jiangsu University of Science and Technology's Basic Research Development Program (No. 2005DX006J).

### REFERENCES

- [1] W. Huntelman, "Automated information system - (AIS) alarm system", in Proc. of the 20th NIST-NCSC National Information Systems Security Conference, 1997, pp. 394-405.
- [2] S. Staniford-Chen, S. Cheung, R. Crawford, et al., "GrIDS: a graph based intrusion detection system for large networks", in Proc. of the 19th National Information System s Security Conference, Vol. 1. National Institute of Standards and Technology, 1996, pp. 361-370
- [3] P.A. Porras and P.G. Neumann, "EMERALD: event monitoring enabling responses to anomalous live disturbances", in Proc. of the 20th National Information Systems Security Conference, National Institute of Standards and Technology, 1997, p. 13.
- [4] E.H. Spafford, "Intrusion detection using autonomous agent", Computer Networks, 2000, 3(4): 547-570.
- [5] D. Dasgupta and H. Brian, "Mobile security agent for network traffic analysis", in Proc. of DARPA Information Survivability Conference and Exposition II (DISCEX-II), June 2001, Anaheim, CA, pp. 332-340.
- [6] W. Jansen, P. Mell, T. Karygiannis, and D. Marks, "Mobile agents in intrusion detection and response", in Proc. of the 12th Annual Canadian Information Technology Security Symposium, Ottawa, Canada, June 2000, p. 12.
- [7] S.A. Hofmeyr, S. Forrest, and A. Somayaji, "Intrusion detection using sequences of system calls", in Journal of Computer Security, Vol. 6, 1998, pp. 151-180.
- [8] J. Kim and P. Bentley. "Towards an artificial immune system for network intrusion detection: an investigation of dynamic clonal selection", in Proc. of the Congress on Evolutionary Computation, Honolulu, USA, 2002, pp. 1015-1020.
- [9] J. Kim, P. Bentley, U. Aickelin, et al., "Immune system approaches to intrusion detection- a review", Natural Computing, 2007, 6: 413-466.
- [10] U. Aickelin, J. Greensmith, and J. Twycross, "Immune system approaches to intrusion detection - a review", in Proc. of ICARIS, 3rd international Conference on Artificial Immune Systems, LNCS 3239, Springer-Verlag, Catania, Italy, 2004, pp. 316-329.
- [11] M. Glickman, J. Balthrop, and S. Forrest, "A machine learning evaluation of an artificial immune system", Evolutionary Computation, 2005, 13(2): 179-212.
- [12] J. Gomez, F. Gonzalez, and Dasgupta D, "An immune-fuzzy approach to anomaly detection", in: Proc. of the 12th IEEE International Conference on Fuzzy Systems (FUZZIEEE), Vol. 2, May 2003, pp. 1219-1224.
- [13] A. Zainal, M.A. Maarof, and S.M.Shamduddin, "Feature selection using rough set in intrusion detection", in Proc. IEEE TENCON, 2006, p. 4.
- [14] B.j. Kim and I.k. Kim, "Kernel based intrusion detection system", in Proc. IEEE ICIS, 2005, p. 6.G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529-551, April 1955