An Automatic Co-stimulation Algorithm for LAN Artificial Immune Systems

ZHAO Tie-Shan, LI Zeng-Zhi
Graduate School of Computer Architecture and Network
Xi'an Jiaotong University
Xi'an, China
zts6389_cn@sohu.com

MAO Wan-Biao, ZHU Jia-Jun
Xichang Satellite Launching Center

Xichang, China
zts6389_cn@sina.com.cn

***Abstract*—Co-stimulations are very important in LAN AISs, but there are few automatic co-stimulation algorithms or approaches at present, and almost all of co-stimulations are implemented by network-administrators. We try to find an automatic co-stimulation algorithm to replace the administrators in LAN AISs. Detectors employed to detect abnormal packets are called NS-detectors. Detectors employed to detect normal packets are called PS-detectors. Each NS-detector can detect a different kind of potential abnormal packets. Each PS-detector can describe a different kind of normal network behaviors. The NS-detector generation algorithm and the PS-detector generation algorithm are presented first. Based on them, an automatic co-stimulation algorithm for LAN AISs is described in detail. Co-stimulations in LAN AISs can be implemented automatically with the algorithm. The administrators can still implement co-stimulations, but their co-stimulations are not important. The feasibility of the algorithm is illustrated with our experiments in laboratory. Our further work is to illustrate more the feasibility of the algorithm in the Internet environment.**

***Keywords*—co-stimulation, algorithm, automatic, PS-detector, NS-detector**

I. INTRODUCTION

LANs are applied widely in schools, enterprises, and so on. They help people a lot in their daily learning and working, and their security is commonly concerned. In 1974 [1], network theory was tried to explain the biological immune system. From then on [2], more and more artificial immune theories have been researched, and all kinds of artificial immune systems (AISs) have been built for LAN anomaly intrusion detection.

In LAN AISs [3], when a mature detector or a memory detector detects a potential abnormal data packet, a co-stimulation is needed to confirm whether the packet indicates a real intrusion or not. If the co-stimulation confirms that the packet indicates a real intrusion, the mature detector evolves into a memory detector. Otherwise, the mature detector or the memory detector is discarded, and new mature detectors are generated. So, co-stimulations are very important in AISs; they confirm whether a real intrusion occurs or not and decide whether mature or memory detectors are discarded or not.

But at present [2,3,4,5], almost all of co-stimulations in LAN AISs are implemented by network-administrators. There are few automatic co-stimulation algorithms or approaches can be employed. The administrators play vital roles in co-stimulations. If the administrators are versed in intrusion detection and are conscientious, co-stimulations will be implemented correctly. Otherwise, co-stimulations may not be done correctly. Because the administrators are so important in co-stimulations and not all of them are employable, in this paper, we try to find an automatic co-stimulation algorithm to replace the administrators in LAN AISs and to improve the adaptabilities of LAN AISs.

Our algorithm is based on following assumptions. Detectors employed to detect abnormal packets are called NS-detectors, and detectors employed to detect normal packets are called PS-detectors. It is assumed that there are some common features in some abnormal network behaviors. This assumption is reasonable; it has been certified by lots of AISs; and it is the theoretical foundation that less NS-detectors can be employed to detect much more abnormal packets. Analogously, it can be assumed that there are some common features in some normal network behaviors, and less PS-detectors can be employed to detect much more normal packets. Therefore, NS-detectors can be taken as the description of potential abnormal network behaviors, and PS-detectors can be taken as the description of normal network behaviors.

Based on previous assumptions, when a NS-detector detects a potential abnormal packet, a co-stimulation can first be implemented automatically by other PS-detectors and NS-detectors; and then an administrators' co-stimulation is asked for in a given time. If the administrator doesn't implement a co-stimulation in the given time, the potential abnormal packet will be confirmed normal or abnormal according to the automatic co-stimulation algorithm.

In section 2, the automatic co-stimulation algorithm is described in detail. Some preliminary experimental results are given in section 3, and a few conclusions and future work are presented in section 4.

II. THE AUTOMATIC CO-STIMULATION ALGORITHM FOR LAN AISS

In a LAN, data packets are broadcasted to every host, and each host can capture every packet in the LAN.

They are assumed that there are *m* intrusion detection hosts in a LAN; each host has its own self-set, NS-detector set and

PS-detector set; all hosts produce their self-sets, NS-detector sets and PS-detector sets asynchronously; each host detects data packets separately; the length of every self or NS-detector or PS-detector is *l*; selves, NS-detectors, and PS-detectors are all binary strings; and the *r*-contiguous bits matching rule is employed to calculate the affinity between two binary strings [6].

A self-set is a collection of binary strings extracted from normal data packets captured by a host in a given time. Because large amount of intrusion traces such as Trojans and viruses hide in packet-bodies (not packet-heads), the binary strings are extracted from packet-bodies.

Based on the negative selection algorithm [6], a NS-detector generation algorithm is presented in this paper as in figure 1. If a randomly generated immature detector doesn't match any self in the self-set and doesn't match any NS-detector in the NS-detector set, it evolves into a mature NS-detector and is added into the NS-detector set; otherwise it is discarded. With this algorithm, in the NS-detector set, a NS-detector may not match another one. So each NS-detector describes a different kind of potential abnormal network behavior and may detect a different kind of abnormal data packets.

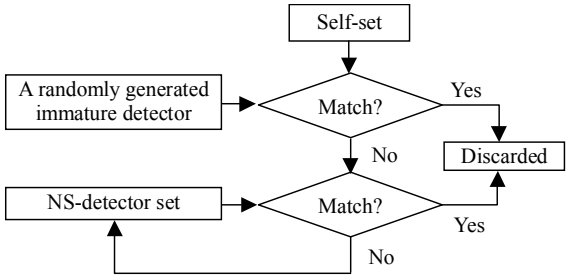

Figure 1. NS-detector generation algorithm

Based on the positive selection algorithm [7], a PS-detector generation algorithm is presented in this paper as in figure 2. If a randomly generated immature detector matches any self in the self-set and doesn't match any PS-detector in the PS-detector set, it evolves into a mature PS-detector and is added into the PS-detector set, and the matched selves are taken away from the self-set; otherwise it is discarded. This process is repeated until there are no selves in the self-set. With this algorithm, in the PS-detector set, one PS-detector may not match another one. Therefore each PS-detector describes a different kind of normal network behavior and can detect a different kind of normal data packets.

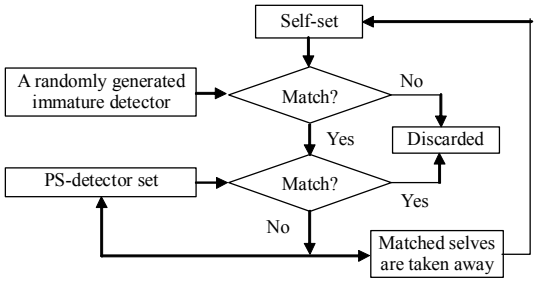

Figure 2. PS-detector generation algorithm

In the dynamical clonal selection algorithm [3], when a NS-detector detects a potential abnormal packet, it asks for an administrator's co-stimulation at once, and the co-stimulation is entirely implemented by the administrator. But in our automatic co-stimulation algorithm, there are 2 co-stimulations. The first is implemented automatically. Only the second is needed to implement by the administrator, and the administrator have to do the second in a given time. If the administrator doesn't do the second in the given time, the packet is confirmed abnormal or normal by the first.

See figure 3. When a NS-detector detects a potential abnormal data packet, it broadcasts the packet to all other NS-detectors and PS-detectors in the LAN, and the affinity between every NS-detector or every PS-detector and the packet is calculated with the *r*-contiguous bits matching rule [6]. Both the number of NS-detectors and the number of PS-detectors who match the packet are counted up, and they are N_1 and N_2 respectively.

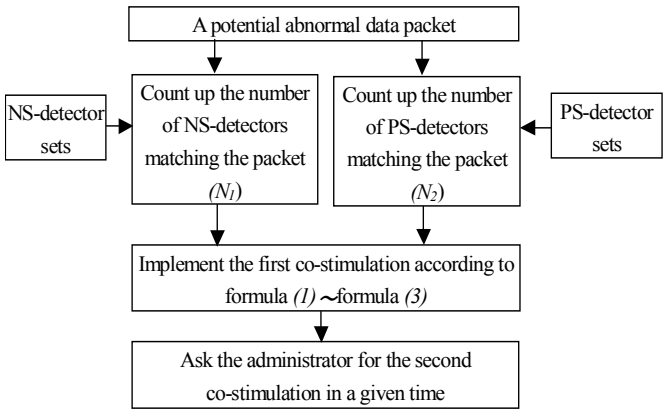

Figure 3. Automatic co-stimulation algorithm

A simple automatic co-stimulation approach may be as follows. If $N_1 \geqq N_2$, the packet is confirmed abnormal, otherwise it is confirmed normal. But because the probability of a NS-detector detecting a potential abnormal packet and the packet being really abnormal may be different from the probability of a PS-detector detecting a potential abnormal packet and the packet being really normal, a better automatic co-stimulation approach should embody the 2 potentially different probabilities.

They are assumed that if a NS-detector detects a potential abnormal packet, the probability of the packet being really abnormal is P_1; if a PS-detector detects a normal packet, the probability of the packet being really normal is P_2; if N_1 NS-detectors match a packet, the probability of the packet being really abnormal is P_{NS}; if N_2 PS-detectors match a packet, the probability of the packet being really normal is P_{PS}; the co-stimulation function is *Co-stim*. Then following 3 formulas are given.

$$P_{NS} = 1 - (1 - P_1)^{N_1} \quad (1)$$

$$P_{PS} = 1 - (1 - P_2)^{N_2} \quad (2)$$

$$Co-stim=\begin{cases}1 & if \quad P_{NS} \geq P_{PS} \\ 0 & if \quad P_{NS} < P_{PS}\end{cases} \tag{3}$$

Where, P_1 and P_2 are given values by experienced LAN AIS developers. Commonly, because there are much more normal packets than abnormal ones in a LAN, $P_1 \leqslant P_2$. In formula *(3)*, *Co-stim=1* indicates that the packet is abnormal, and *Co-stim=0* indicates that the packet is normal.

If the administrator does not implement the second co-stimulation in a given time *T*, the potential abnormal packet is confirmed abnormal or normal according to formula *(1)* to formula *(3)*. Otherwise, it is confirmed according to the administrator's co-stimulation. When the administrator implements the second co-stimulation, the information of the first co-stimulation such as N_1, N_2, P_{NS}, P_{PS}, and the result of function *CO-stim* will help him a lot. The given time *T* can be several seconds or more or less; it isn't important for the algorithm performance. Although the administrator can implement a co-stimulation as in [3], his co-stimulation isn't important, too.

The approach of mature or memory NS-detectors' evolution or death is the same as that in the dynamical clonal selection algorithm [3]. A PS-detector dies only if it matches a packet confirmed abnormal by co-stimulations.

III. Experimental Results

To validate the feasibility of previous automatic co-stimulation algorithm, some experiments have been done in our laboratory. 5 hosts are linked to a HUB to build a LAN. A host is selected as a broadcaster to broadcast data packets to other hosts. The length of every packet-body is 64 bits, $l=64$, $r=8$, $P_1=0.4$, and $P_2=0.6$. The probability of 2 random binary strings matching at at least *r*-contiguous locations is P_M, then [3]

$$P_M \approx 2^{-r}((l-r)/2+1)=2^{-8}((64-8)/2+1)=0.11328125$$

Each host needs N_R NS-detectors to detect abnormal packets, and the probability that N_R NS-detectors fail to detect an abnormal packet is P_f. Then [3]

$$N_R=ln(P_f)/ln(1-P_M)$$

In our experiments, $P_f=0.1$, then

$$N_R=ln(P_f)/ln(1-P_M)=ln(0.1)/ln(0.88671875) \approx 20$$

The probability that all of the 4 hosts fail to detect an abnormal packet is P_{allf}, then

$$P_{allf}=P_f^4=0.1^4=0.0001$$

Our experiments are as follows.

In stage 1, 5 clean DLL software modules are prepared. Their sizes are 32Kb, 36Kb, 44Kb, 29Kb, and 47Kb respectively. Each of them is disassembled into packet-bodies, and the length of every packet-body is 64 bits. The 4 self-sets are same, and each one is made up of all of the clean packet-bodies. In every self-set, there are not 2 or more elements are same. The clean packet-bodies are broadcasted in the LAN to generate the 4 PS-detector sets and the 4 NS-detector sets. The size of the 4 self-sets, 4 PS-detector sets and 4 NS-detector sets are as in table Ⅰ.

TABLE I. Size of Sets

	Host Ⅰ	Host Ⅱ	Host Ⅲ	Host Ⅳ
Size of Self-set	*22861*	*22861*	*22861*	*22861*
Size of PS-detector Set	*283*	*276*	*294*	*301*
Size of NS-detector Set	*20*	*20*	*20*	*20*

TABLE II. Detection Results with Automatic Co-stimulation

	Module Ⅰ	Module Ⅱ	Module Ⅲ	Module Ⅳ	Module Ⅴ
Alert Times	*129*	*153*	*141*	*151*	*105*

In stage 2, each of the previous 5 DLL software modules is infected with a different real Trojan in the Internet. Their sizes become 43 Kb, 49 Kb, 56 Kb, 42 Kb and 56 Kb respectively. Then the 5 infected modules are broadcasted one by one in the LAN. Without any administrator's intervening, the detection results with the automatic co-stimulation algorithm are as in table Ⅱ. While any of the 5 infected modules is being broadcasted, the automatic co-stimulation algorithm is able to confirm abnormal packets.

Some other similar experiments are done, and their results are analogous to the above.

IV. Conclusions and Further Work

We have described an automatic co-stimulation algorithm, a NS-detector generation algorithm and a PS-detectors algorithm in previous sections. The first algorithm is based on the second algorithm and the third algorithm.

We have illustrated the feasibility of the automatic co-stimulation algorithm with our preliminary experiments. With the automatic co-stimulation algorithm, the administrators can still implement co-stimulations as in [3], but their co-stimulations are not obligatory. Our further work is to generate the PS-detectors and the NS-detectors with the real Internet traffic, and to illustrate more the feasibility of the algorithm in the Internet environment.

The NS-detector generation algorithm is based on the negative selection algorithm [6]. But because a NS-detector may not match another one in the same NS-detector set, each NS-detector can detect a different kind of potential abnormal data packets.

The PS-detector generation algorithm is based on the positive selection algorithm [7]. But because a PS-detector may not match another one in the same PS-detector set, each PS-detector can describe a different kind of normal network behaviors. Our further work on this algorithm is to apply it to normal network behavior descriptions. Compared with short sequences of system calls [8], the PS-detectors are more adaptive, easier to acquire and easier to understand.

References

[1] Jerne N K. *Torwords a Network Theory of the Immune System*. Annual Immunology, vol. 125c, 1974.

[2] Dipankar Dasgupta. *Advances in AISs*. IEEE Computational Intelligence Magazine. Nov. 2006.

[3] J Kim, P Bentley. *Towards an Artificial Immune System for Network Intrusion Detection: An Investigation of Clonal Selection with a Negative Selection Operator*. Proceedings of the Congress on Evolutionary Computation. (CEC), Seoul, Korea, May 2001.

[4] S. Forrest, S. Hofmeyr, and A. Somayaji. *Computer Immunology*. Communications of the ACM Vol. 40, No. 10, pp. 88-96 (1997).

[5] LI Tao. *Computer Immunology*. Beijing: Publishing House of Electronics Industry, July 2004.

[6] S. Forrest, A.S. Perelson, L. Allen, and R. Cherukuri. *Self-nonself discrimination in a computer*. Proceedings of the 1994 IEEE Symposium on Research in Security and Privacy, Los Alamitos, CA: IEEE Computer Society Press, pp. 202-212 (1994).

[7] Seiden P E, Celada F. *A Model for Simulating Cognate Recognition and Research in the Immune System*. J.theor.Biol., 158:329～357,1992.

[8] C. Warrender, S. Forrest, and B. Pearlmutter. *Detecting Intrusions Using System Calls: Alternative Data Models. 1999 IEEE Symposium on Security and Privacy*. pp. 133-145 (1999).