Method for Presenting Danger Signals Based on System Balance

HeYANG[1, a] Si XIONG[1, b]

Computer School, HubeiUniversity of Education, Wuhan 430205, China

[a]email:yanghe@whu.edu.cn, [b]email:xiongsi@gmail.com

Keywords: Artificial Immune Systems, Danger Theory, Danger Signal

Abstract. The second generation Artificial Immune Systems based on Danger Theory are not limited to distinguish Self/Nonself. Instead, the Danger Signals generated when systems are threatened, for this reason, the problems of high training costs and low coverage are effectively avoided. The definition and presentation of danger signals are the most important problems in Danger Theory. most of the current studies set danger signals according to special problems, and the signals are set by human experience. The characteristic of adaptability of the Artificial Immune Systems is destroyed. The function of immune system is to maintain the balance of body, the break of balance would inevitably lead to changes. Take the changes as the possible danger signals are feasible. The theory of differential coefficient is used for defining and calculating danger signals. This method does not rely on prior knowledge of human, it is general and adaptive. A worm is used as the example of malicious software to test this method, and the result proved its validity.

1 Introduction

Artificial Immune Systems (AIS) are Intelligent Systems which Learned from the ideas of the biological immune systems. These systems are used to solve some Specific complex problems.

AIS belongs to the research field of Bio-inspired computationwhich is a branch of computational intelligence, just like neural network and evolutionary computation. At present time, AIS include four branches, they are Negative Selection (NS), Clonal Selection (CS), Artificial Immune Network (AIN) and Danger Theory (DT). Clonal Selection and Artificial Immune Network based on Negative Selection, so these three branches are called “The first generation of AIS” [1]

The first generation of AIS is inspired from Clonal Selection Theory in Biological Immunology. This kind of AIS supposed that the “self” set of immune system is known, from negative selection, a lymphocyte set which can detect unknown “nonself” is obtained. Because of the characteristic of adaptive unknown nonself detection, the first generation of AIS is used in anomaly detection, data mining and relative fields. However, in the practical application self and nonself sets are massive, high training cost and low coverage are the main problems. With the deepening of AIS research, these problems are more prominent, and difficult to be solved in existing models[2].

In 2002, Uwe aickelin in university of Nottingham presented an AIS model inspired from the Danger Model in Biological Immunology. [3] His model is called Danger Theory in AIS. Danger Model is a hypothesis in immunology. In this hypothesis, danger signals released by necrotic cells which are suffered from pathological changes or pathogens’ invasion are the keys for starting immune responses, and nonself is no longer the most important defense object. Therefore, the DT model take finding "danger signal" as the core. This model provides a new way of thinking to reduce the cost of training and improve the identification rate. So the Artificial Immune systems based on this model is called the second generation of AIS[1].

InBiological Immunology danger theory, danger signalsare the keyfor immune response. In order to design the model of Artificial Immune Systems based on Danger Theory, firstly, we need to define danger signals. Team in university of Nottingham presented the DCA （DCA ,Dendritic Cell Algorithm） model[4] and the artificial innate immune systems experiment platform Libtissue[5]. In their models, signals are divided into danger signals, safe signals and PAMPs signals. All these signals are manually set for a certain problem based on empirical knowledge, does not comply the adaptive characteristics of AIS. A. Iqbal concerned the danger susceptible codon, based on this

proposed presented a model for artificial APCs (Antigen Preseting Cells) called DASTON (Danger Susceptible Data Codon), and verified in the system call sequence recognition experiment. However, he did not specify what is the danger signal[6].

From the viewpoint of balance, the function of the immune system is to ensure the body's balance. Immune response is to adjust the imbalance, and rebalance it. In general, a stationary system does not change the balance， if the imbalance happens, there is sure to be some changes destroied the balance, and to the immune systems, the changes are the danger signals .

In this paper, the function of artificial immune system is to maintain the balance of the protected system.Various system changes are danger signals which break the balance, and they are the goal for immune response. Study danger signals, in essence, is to find the changes, and extract changes, figure out how to start immune response from changes.

Derivative and differential are instruments for describing changes.In this paper we drawing the general definition and calculating method of danger signals based on derivative and differential, a group of examples demonstrate the feasibility.

2 Background

2.1 Biological principles of danger signals

Danger Theory of biological immunology is presented by Polly.Matzinger in 1994. In this theory, which shoud be defensed by immune systems are not nonselves but potential danger[7]. Immune responses are caused by danger signals.Danger signalsare the core issue of the danger theory.

The current study shows that, the danger signals are released material such as Hsps (Heat Shock Proteins), S100 proteins, uric acid and so on, when death of the cell via necrosis. The material is also called Danger-associated molecular patterns. So danger signals are endogenous signals, they are releaseby the body itself in the case of damage to the body.

Antigen Presenting Cells are an important kind of cells in biological immune systems, which link the innate immune system and the adaptive immune system. Theyare generally found in tissues, take the role of guards. When the concentration of a danger signals reaches a certain level, APCs are activated. Actived APCs captureantigens, leave the organization into the lymph nodes, release costimulatory signals to active lymphocytes and present the capturedantigens to lymphocytes, eventually cause an immune response. In the view of the danger theory, the whole process of immune response is shown in Figure 1.

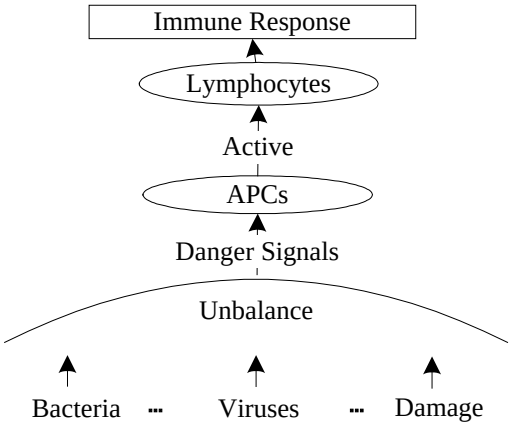

Fig1. Process of immune response in Danger Theory

Danger signalshave the following two features：

- danger signalsare endogenoussignals, that is to say, it is released by the damaged cellsthemselves.
- danger signals are released when the balance is broken and death of the cells via necrosis happen, Under normal conditions does not appear.

2.2 Inspiration of differential

Changes of the computer systems, are the display of unbalance. They are endogenous signals, consistent with characteristic of danger signals. Extraction system changes, as danger signals, in line with the biological principles of danger signals.

Mathematics is the foundation of natural sciences. It describes the common law of natural sciences. whether biology or computer science followds the fundamental mathematical principles. In calculus the speed of changes is defined to be derivative and the degree of changes is defined to be differential . Ispired from calculus, the changes of the state of a computer system indicate the possible dangers, and the trend of changes indicates the degree of dangers. Since the method which is used to describe changes in mathematics is called differential coefficient. Considering the computer data are discrete, numerical differentiation is used to describe the changes.

In numerical analysis, numerical differentiation describes algorithms for estimating the derivative of a mathematical function or function subroutine using values of the function and perhaps other knowledge about the function.

f(x) is defined in [a,b], $a \leq x_0 < x_1 < \cdots x_n \leq b$aren given nodes in [a,b]. The values of function in the above-mentioned nodes are $f(x_0), f(x_1), \cdots, f(x_n)$. Approximation of$f, f', f''$, can be calculated.

$$f'(x_i) \approx \frac{f(x_i) - f(x_{i-1})}{h} \quad (1)$$

$$df(x_i) \approx f(x_i) - f(x_{i-1}) \quad (2)$$

3 The Expression of Danger Signals Based on Changes

Firstly, present the definition and description of relevant concepts. Based on this proposed the general definition of danger signals.

3.1 Definition of related conception

Definition1 S(System) A system is a set of interacting or interdependent components forming an integrated whole[9].

Definition2 SS(System States) System status refers to a system snapshot at a particular moment. It is a function of many system variablesSS=f(V). V is a collection of system variables, V={ vi| i ∈ N}, vi is one of the system variable.

Take computer system as an example, the system variables include memory usage, CPU usage, upstream network traffic, downstream network traffic, stream of TCP, stream of UDP and so on.

Definition3 EN(Environment) the surroundings of a physical system that may interact with the system by exchanging mass, energy, or other properties.

Definition4 EV(Events) Internal or external factors which cause the system state change.

Definition5 BS(Balance State) IF the system status does change over time, the system is said to be in balance. balance state of the system is relatively static state.

Definition6 IS(Imbalance State) Under the action of event EV, the state of the system changes over time, the system is Imbalance. Generally if there is a danger, the system state transit from balance to imbalance.

F is a mapping of the system from balance to imbalance, F (BS, EV) = IS.When system state is inbalance state, it can be changed to imbalance state by the event EV.F' is a mapping of the system from imbalance to balance, F'(IS, EV')=BS.When system state is inimbalance state, it can be changed to balance state by the event EV'.

Definition7 SC(System Changes) V is the set of system variables, R is the frame of reference for system status, the changes of system states

$$\mathrm{SC} = \mathrm{dSS} = \frac{df(V)}{dV} = \frac{\partial SS}{\partial v_1}\Delta v_1 + \frac{\partial SS}{\partial v_2}\Delta v_2 + \cdots + \frac{\partial SS}{\partial v_n}\Delta v_n \quad (3)$$

the set of system

variables $\mathrm{V} = \mathrm{G}(\mathrm{R}) = \{g_1(R), g_2(R), \cdots g_n(R)\}$,$v_1 = g_1(R), v_2 = g_2(R), \cdots, v_n = g_n(R)$

Changes of System variable V o the frame of reference R can be expressed as

$$\frac{dV}{dR} = \frac{d\{v_1, v_2 \cdots v_n\}}{dR} = \left\{\frac{dv_1}{dR}, \frac{dv_2}{dR}, \cdots, \frac{dv_n}{dR}\right\} \quad (4)$$

The frame of reference can be time, or may be event, the amount of data, etc

3.2 Definition of danger signals

Danger signals come from several kinds of system variables, all the danger signals make up a set. The definition of danger signals is showed as definition 8.

Definition8 DS(Danger Signals) Danger signals are changes of system variables related with system imbalance. DS ={ dsi| i ∈ N}, it is a subset of system variables changes.

In this paper, all the system variables changes are taken as potential danger signals, $DS = dV$. dsi is a single signal. Artificial APCs act as filters for removing system variables unrelated to changes. Artifical APCs and the receptors can evolve by themselves, With the evolution of generations, receptors which can match danger unrelated changes will be nature selected.

3.3 Expression of danger signals

Combined with the definition of danger signals, the expression of danger signals are presented as follow. R is the frame of reference.

$$\mathrm{DS} = \mathrm{dV} = \{dv_1, dv_2 \cdots dv_n\} = \{dg_1(\mathrm{R}), dg_2(\mathrm{R}), \cdots dg_n(\mathrm{R}),\} \quad (5)$$

Referring to the method of numerical differentiation, Value of a single danger signal and set of danger signals are expressed as follows.

$$ds_i \approx g_i(R_i) - g_i(R_{i-1}) \quad (6)$$

$$\mathrm{DS} = \mathrm{dV} = \{dv_1, dv_2 \cdots dv_n\} = \{dg_1(R), dg_2(R), \cdots dg_n(R)\} \quad (7)$$

3.4 Structure of danger signals

Danger signals come from changes system variables. In order to distinguish different system variables and corresponding danger signals, there are several different kinds of Toll-like receptors in APCs, each kind of TLRs (Toll-like receptors) correspond to a certain kind of danger signals. The structure of danger signals is designed as Figure2.

ds_Category	ds_Name	ds_Value

Fig2. Structure of danger signals

ds_Category corresponding to the type of danger signals, preliminary analysis suggests thatds_Category={ Gradient, Jump, Mutant, Mixed}. The ds_Category decides which threshold should be choose when TLRs matches danger signals.

ds_Name corresponding to the name of danger signal. It shows us which kind of system variables should the danger signal correspond to

$$\mathrm{ds_{Name}} = \{v_1, v_1, \cdots, v_n\}$$

Take computer system for an example, $\mathrm{ds_{Name}} = \{CPU, Memory, \cdots\}$. ds_Name uniquely determines which kindof TLRs receptors match this kind of danger signals.

ds_Value corresponding to the values of danger signals,that is, the amount of change. Ds_Value matches the threshold of TLRs, and decides whether the danger signals should be concider to be effect.

3.5 Fusion of danger signals

In biological immune systems, APCs are the middleware for fusing danger signals and activing Lymphocytes. Inspired by the works of APCs, we can build artificial APCs as fusion device for danger signals.

The strength of costimulatory signal is determined by the size of the active population of APCs. Costimulatory signal is a sign of a dangerous state of the system, it is also the alarm signal.

The artificial APCs Population structure is designed as figure3.

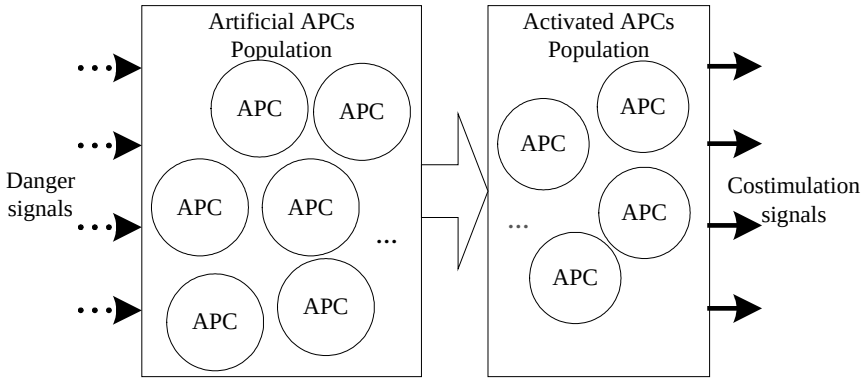

Fig3. Fusion mode of danger signals

The initial population size is P, if the activated APCs population over P', the concentration of Costimulatory signal is enough to activated lymphocytes.

4Experiments and Results Analysis

Worm.Win32.Smelles is chosen as an example of malicious software for verifing the effectiveness of the danger signal definition based on differential. 12 computer system variables are chosen. The monitoring system variables are all randomly selected as table 1. Time is chosen as the frame of reference. The sampling period T = 1 second, the sliding window W = 5.

Tab1. Selected System Variables

variables	Discription of variables	variables	Discription of variables
CPU	CPU Usage	ICMP	ICMP packets traffic
Run	Registry startup item 1	IP	IP packets traffic
Runonce	Registry startup item 2	TCP	TCP packets traffic
File	The number of file operationAPI calls	TCPFIN	TCPFIN packets traffic
Keylogger	The number of keyboard operationAPI calls	TCPPSH	TCPPSH packets traffic
Socket	The number of Communication API calls	ARP	ARP packets traffic

For comparison, the experiment is divided into two groups. Reference group is collected during normal operation of the system, Malware group is collected when the worm is running. The population of APCs P=150, there are 5 TLRs receptors in one APC, when the activated APCs over 2/3 of the total population, alarm.

Result of the experiment shows that in malware implanted group, the alarm is significantly higher than the reference group.Although the reference group also produces a small amount of alarms, however, the alarmsare sporadic, random, it is more likely caused by experimental environmental noise. Alarms in the malware group are concentrated and regular, it is more likely caused by malware.Experimental results shows that the danger signals define method based on balance is effective and adaptive.

5 Conclusion

Function of the immune system is to maintain system balance, and find changeswhich cause imbalance. These changes are danger signals for immune response.

From the perspective of system balance, this paper descripted changes inspired fromthe law of differential, presented the definition and description of danger signals. Considering the discreteness

of computer data, danger signals are calculated by the method based on numerical differentiation. Artificial APCs population is built to fuse danger signals and determine the state of the system.

Preliminary experiments confirmed the validity and adaptability of the method.

6 Acknowledgment

This work is supported by the natural science foundation of Hubei ProvinceNo. 2014CFB569. This work is also supported by the research project of Hubei Province Department of Education Grant No.Q20133008.

References

[1] GREENSMITH J, WHITBROOK A, AICKELIN U. Artificial Immune Systems [M]. Handbook of Metaheuristics. Springer. 2010: 421-48.

[2] GREENSMITH J, AICKELIN U, TEDESCO G. Information fusion for anomaly detection with the dendritic cell algorithm [J]. Information Fusion, 2010, 11(1): 21-34.

[3] AICKELIN U, CAYZER S. The danger theory and its application to artificial immune systems; proceedings of The 1st International Conference on Artificial Immune Systems (ICARIS 2002), Canterbury, UK 2002 [C]. Citeseer.

[4] GREENSMITH J, AICKELIN U. The Deterministic Dendritic Cell Algorithm [M]. 2008: 291-302.

[5] TWYCROSS J, AICKELIN U. Libtissue - Implementing innate immunity; proceedings of 2006 IEEE Congress on Evolutionary Computation, CEC 2006, July 16, 2006 - July 21, 2006, Vancouver, BC, Canada, 2006 [C]. Inst. of Elec. and Elec. Eng. Computer Society.

[6] IQBAL A, MAAROF M A. Towards Danger Theory Based Artificial APC Model: Novel Metaphor for Danger Susceptible Data Codons [M]. 2004: 161-74.

[7] MATZINGER P. Tolerance, danger, and the extended family [J]. Annual Review of Immunology, 1994, 12(1): 991-1045.

[8] MATZINGER P. The danger model: A renewed sense of self [J]. Science, 2002, 296(5566): 301-5.

[9] http://www.merriam-webster.com/dictionary/system