ISSN NO: 2770-4491
Date of Publication:07-11-2022

A Controllable and Adaptable Computer Virus Detection Model

***Abdumuminov Abdurafiq Abdurashidovich **Ibragimov Jalaliddin Obidjon o'g'li**
***** Shoraimov Khusanboy Uktamboyevich**
*Republican center for management of telecommunications networks of Uzbekistan. SUE.
** Teacher of the Department, "Systematic and Practical Programming", Tashkent University of Information Technologies named after Muhammad Al-Khwarizmi, UZBEKISTAN
*** Teacher of the Department, "Systematic and Practical Programming", Tashkent University of Information Technologies named after Muhammad Al-Khwarizmi, UZBEKISTAN

***Abstract*—**This paper presents a controllable and adaptable computer virus detection model based on immune system, improves the immune mechanism of the model by introducing the thought of variable detector radius, and achieves the detector radius be set and adjusted automatically. The virus detection model has some effects on solving the conflict between the number of detectors and the scope of detection. Thus, it enables the system to have better dynamic and self-adaptability.

Keywords*-Artificial Immune; detector; controllable adaptive; virus detection*

I. Introduction

The problem of biological immune system (BIS: Biological Immune System) and detecting virus in computers is almost the same. one of the most natural applications in the immune calculation is virus detection. The principle of immunity has been applied to the virus detection in the Forrest and the virus laboratory at IBM. After that the research in Computer Virus Immune System(CVIS) has become more and more popular. People like S. Homeyr, J. O. Kphart, P. Harmer, and T. Okamoto have put forward a series of computer-virus-immunity model. However, these existing CVIS has not been widely used, mainly because these CVISs still have some shortcomings. One of the most important shortcomings is the contradiction between system's efficiency, detection rate and error rate (error-affirmative and error-negative). Because library of CVIS itself is too large, the number of detectors is too large. And the cost of training mature detectors and the size of the monitors' set shows an exponential relationship, which directly resulted in time inefficiencies. But if we decrease the size of the monitors' set, the system will drop its detection rate and raise its error-affirmative rate. For example, in the virus laboratory at IBM's immune system, because the tolerance- transaction takes too much time, we have to change our original mind that want to do distributed defense to gather the tolerance-transaction to treating and training on a fixed, which make the system effective and feasible. But in fact, this method delays the response-time to detecting unknown viruses at client. What's more is that if the analytical center collapses or the analytical center loses contact with the client, the client will not be able to defense against unknown viruses. So, in this paper, a controllable and adaptable computer virus detection model based on immune system, AISVDM (Virus Detection Model base-on Artificial Immune System), has been put forward. It'll discuss with solving the conflict between the number of detector and the scope of detection and the model's dynamic and self-adaptability.

II. Immune-System-Based Virus Detection

In an immune system, the most important question is how to distinguish non-autologous organism from the body. Because the information in computer software system will be eventually reduced to a binary string, so in fact, the computer virus detection is a problem which is according to some certain rules and a priori knowledge analyzing and detecting problems in binary strings.

A. The definition of immune components

ISSN NO: 2770-4491
Date of Publication:07-11-2022

In the natural immune system, the body represents a set of all normal cells in biological body. And non-autologous body is on behalf of a set of external cell. In this model, the definition of the state space of computer virus detection is that using a fixed-length binary string consisting of finite set U to express all of the cells, and (a binary string is the same as a hexadecimal string), i is natural number. U can be divided into two subsets: N expresses non-autologous body; S represents the body. The relationship between the two is as. D, the definition of a set of immune cells (detector), is as D ={ d|d=< a, age, count, max-age>, a U JI age, count, max- age Z+ }, a for antibody, age for detector' s age, count for the number of matching the antigens, that's called accumulative affinity, Z+ for the positive integer set, max-age for the largest age of detector. Here the length of a check for 24. It means the definition of antibody is a length-of-24 hexadecimal virus signature.

B. Implementation of the immune mechanism 1) Negative selection mechanism Immune mechanisms which are mainly used in the system negative selection included negative selection mechanism and clonal selection mechanism. Negative selection process is the processes of immune cells' selftolerance to ensure not identify immune cells as viruses in the detection. In the process of self-tolerance, set of immune cells will eliminate all detecting cells that can identify autologous cells in immune cells' set. This process can be used like following formula:

$$\text{ftolerance}(D)=D-\{d \mid d \ B \ y \ \text{Self}(\text{fmatch}(d.a, y)=1)\} \quad (2)$$

The merits of the algorithm are simple and easy to realize and to analyze the problem space and abnormal detection without a priori knowledge. It has a strong robustness, parallel, distributed detection and so on. However, there are some weaknesses like: detectors are produced by random algorithms, it is inevitable that a lot of invalid detector will be produced; and its time complexity grew exponentially with the size of self-set, it is difficult to meet the complex system of timeliness requirements; in addition, nonself space can not be covered completely by the detectors. Therefore, the question in based on negative selection methods is how to cover nonself space by detectors and the number of detectors, that is, how to use less detectors to cover nonself space. In the literature, De Castro proposed a simulated annealing algorithm method of induction of group diversity. Drawing on the idea, the simulated annealing method will be used to optimize the detector to ensure hardly cover the self-space in premise; expand the set of detector coverage not my space; make sure that in the process the number of detectors are fixed. Because of wanting to expand the coverage of nonself space, the first objective function is the detector volume, just like formula (2), the constraint condition is that the detector can't fall into self-space.

$$C\ D = \text{Volume}(D) \qquad s.t. D \not\subset \text{self} \qquad (3)$$

C(D)is for objective function Volume(D) is for detectors space.

However, calculating the detector's volume wastes too much time. So it will be changed to an expression that can be calculated easily and reflect the initial purpose. Intuitively, expanding on nonself spaces detector coverage will reduce the overlapping between detectors. So the maximum optimization problem can be converted into minimum optimization problem, added constraints to the objective function at the same time.

Choosing: If the candidate-detector has identified any one element of self-set, this detector should be eliminated. Else put this detector inside the set of detectors.

• Using optimization algorithms of set of detectors to optimize mature detector, then get D' (the number of detectors is N').

• If _ N'=N'-N<0, N detectors will be generated randomly, and using step to check, or the initialization is done.

Set D consist of 1000 randomly generated detectors; the initial value of adjacent radius of detector rnear is two third of each detector length, that means, rnear= 16; the initial parameter t =100; iterative times for each t L =10; evolutionary generations nummax = 200; attenuation rate of adjacent radius _ = 0.8; attenuation coefficient of temperature _ = 0.8. Then optimize set of detectors by using simulated annealing algorithm. The overlapping between detectors mutate with evolutionary generations as Figure 1. It can be seen from the figure, with evolutionary generations increasing. The overlapping between detectors decrease and the curve decline rapidly at the initial stage. It proves that using simulated annealing algorithm can decrease the overlapping between detectors. Finally, using negative selection algorithm filters detectors and deletes detector which is in the self-set.

ISSN NO: 2770-4491
Date of Publication:07-11-2022

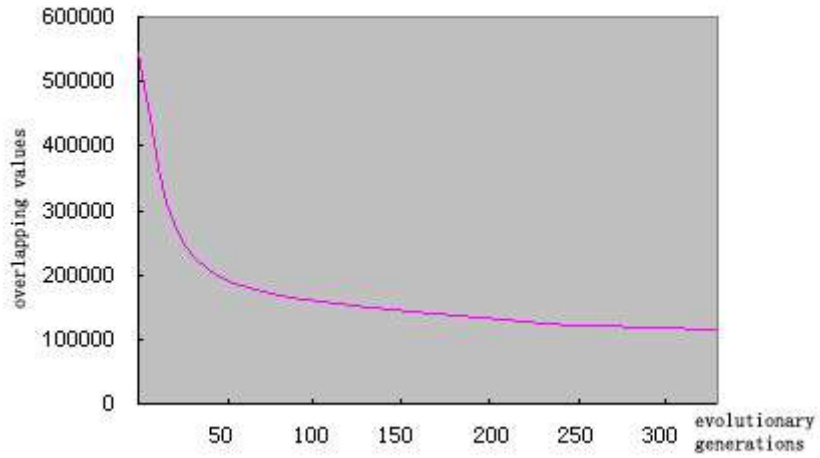

Figure 1 The overlapping between detectors mutate with evolutionary generations

II. Conclusions

Biological immune system becomes a complete and powerful immune system through natural evolution. At present, the research of biological immune mechanism is still imperfect. Relatively speaking, the research of computer-imitation on biological immune has just started. It's the same as the computer immune mechanism. Furthermore, the biological immune system is a highly parallel system, but computer virus immune system can't be parallel as biological immune system. So the most exigent problem which can be applied and have practicality is improving the efficiency of system in the situation of limiting resources and having a good time efficiency, making system have nice dynamics, controllability, self-adaptability and robustness. In this paper, these problems have been discussing and built artificial immune model AISVDM (Virus Detection Model base-on Artificial Immune System) for detecting viruses. In this model, the concept of variable radius of the detector has been introduced and drawing on simulated annealing algorithm for optimization as well as automatically setting and adjusting of radius of detector. It has effect on solving the conflict between the number of detector and the scope of detection. Using optimized dynamic clonal selection to study and detect viruses make the system has better dynamic and self-adaptability. But there are a lot of questions need to be further explored.

1) In theory, it can detect any unknown viruses by CVIS inheriting the thought of distinguish self or nonself in BIS. But how can we ensure the self-information is safe, how to kill the virus and to recover the system faster and more efficient? Whether a more direct and effective way to do that exist? And can it be found through reaching BIS?

2) The model set up to avoid a genetic control, cell signaling and other complex questions on biological immune system. How to put these mechanisms into the model and what will happen to dynamic, adaptability, security and time efficiency of the system?

3) The artificial immune system algorithm is just for specific cases. Profound and significant research achievement about the algorithm or even the calculation and estimation, proving the convergence of whole system is still a fat lot.

References

1. Stephanie Forrest, Alan S.Perelson, Lawrence Allen, Rajesh Cherukuri. Self-Nonself Discrimination in a Computer[C]. In Proceedings of IEEE Symposium on Computer Security and Privacy, 2019
2. Kephart J O. Biological inspired defenses against computer viruses[A]. Proceeding of 14th Joint Conference on Artificial Intelligence[C]. Los Alamitos: IEEE Computer Society Press,2019.
3. J.O. Kephart, Gregory B. Sorkin, Morton Swimmer, and Steve R. White. Blueprint for a Computer Immune System[C]. Proceedings of the 1997 International Virus Bulletin Conference, San Francisco, California, October, 2016.
4. De Castro L N.& Von Zuben F J.Artificial Immune systems Part -Basic Theory and application[J].Technical Report-RT DCA 2017