An Immune-Based Model for Computer Virus Detection

Tao Li[1], Xiaojie Liu[1], and Hongbin Li[2]

[1] Department of Computer Science,
Sichuan University, Chengdu 610065, China
litao@scu.edu.cn
[2] Department of Electrical and Computer Engineering,
Stevens Institute of Technology,Hoboken, NJ07030, USA
hli@stevens.edu

Abstract. Inspired by biological immune systems, a new immune-based model for computer virus detection is proposed in this paper. Quantitative description of the model is given. A dynamic evolution model for self/nonself description is presented, which reduces the size of self set. Furthermore, an evolutive gene library is introduced to improve the generating efficiency of mature detectors, reducing the system time spending, false-negative and false-positive rates. Experiments show that this model has better time efficiency and detecting ability than the classical model ARTIS.

1 Introduction

As the fast development of Internet, the generating and spreading speed of new computer viruses is getting higher and higher. Then, computer viruses and worms are becoming an increasing problem in the world [1,2]. Therefore, it is necessary to detect and eliminate computer viruses, especially the unknown viruses, in real-time. However, it is very difficult for traditional preventing methods [3-5] to solve this problem effectively. In recent years, researchers have taken some researches on the computer network topologies and the spreading mechanism of computer viruses [6-8], then presented some methods to restrain virus spreading [9-11]. These methods can reduce the speed of virus spreading, however, they can not prevent virus spreading [11]. Especially, the problem for unknown virus detection is still not solved.

The problems found in computer security systems are quite similar to the ones encountered in Biological Immune Systems (BIS). BIS has successfully solved the problem of unknown virus detection [12]. Therefore, Artificial Immune System (AIS) [13-15] is considered as a new way to defeat fast-proliferating computer viruses. In 1994, Forrest presented a method of computer virus detection based on the negative selection algorithm[16], which is the first time to use immune mechanism for virus detecting and has greatly promoted the research of computer virus immune system (CVIS). The most important works should be the general

Y.G. Desmedt et al. (Eds.): CANS 2005, LNCS 3810, pp. 59–71, 2005.

framework ARTIS for AIS and the computer virus immune model proposed, respectively, by Hofmeyr [17,18] and Kephart [19,20]. In ARTIS, the concepts and mechanisms of BIS, including self, nonself, self tolerance, immune cell (detectors), memory cell (memory detectors), and costimulation were well simulated. Many CVISs are mainly derived from ARTIS. For example, the computer virus detection system proposed by Okamoto and Ishida [21], the agent based computer virus immune architecture proposed by Harmer [22], and the HMM [23] based computer immune model proposed by Jensen [24]. Different from ARTIS, the computer virus immune model [19,20] proposed by IBM laboratory uses only partial immune mechanisms, however, some other techniques such as automatic extraction of computer virus signatures [19], virus trap [20], etc. have also been adopted.

There are three major defects in the present CVISs: The first is that the self set is very large in size. For example, during the experiments of LISYS [25], a famous application of CVIS based on ARTIS, Hofmeyr and his colleagues collected over 2.3 million self elements in 50 days.The cost for mature detector training is exponentially related to the size of self set [22], making it impossible to directly collect self data from the network for the self tolerance of immature detectors. LISYS has to aim at the detection of 7 kinds of network intrusions, where the services provided by the network, as well as the normal network activities, were simplified in order to decrease the size of self set. After laborious and complicated classification, Hofmeyr finally selected over 3900 elements as self for the tolerance process of the detectors, reducing the training cost for the tolerance of detectors. However, the computation cost is still high.

The second deficiency is that the definitions of self and nonself in the system are described in a static way with almost no changes. However, it is very difficult to use a fixed definition for self and nonself in most practical applications. Furthermore, the roles of self and nonself may exchange at times, e.g., the legal network behaviors today may be dangerous tomorrow, and vice versa. Therefore, it is necessary to update the definitions of self and nonself from time to time. The static description model for self/nonself lacks the adaptability, and thus cannot cater for the network monitoring in the real network environment.

The third, the absence of rigorous quantitative descriptions in most presented CVIS models results in the randomicity of CVIS implementation. Therefore, it is not convenient to put these models into practical applications.

The above three problems have become the major obstacles to CVIS applications. Inspired by biological immune systems, a new immune-based model for computer virus detection is proposed in this paper. Quantitative description of the model is given. A dynamic evolution model for self/nonself description is presented, which reduces the size of self set. Furthermore, an evolutive gene library is introduced to improve the generating efficiency of mature detectors, reducing the system time spending, false-negative and false-positive rates. Experiments show that this model has better time efficiency and detecting ability than the classic model ARTIS.