

## 计算机技术

# 基于生物免疫突变的入侵容忍触发模型

雷丽萍

(渭南市委党校,渭南 714000)

**摘要** 入侵检测技术是保证网络安全的核心技术之一。智能技术是其中一个重要的分支,而且入侵检测的智能化也是入侵检测技术的一个重要发展方向。但是目前的入侵检测系统只是差强人意。因此,以记忆原理为出发点,将免疫原理、突变理论等融入到当前的入侵检测系统框架中,提出了一种基于生物免疫突变的入侵容忍触发模型,试图在如何使得入侵检测系统更加智能化的发展方向上提供一定的参考和提示。

**关键词** 记忆原理 入侵检测 容忍突变 免疫原理

中图法分类号 TP183; 文献标志码 A

## 1 入侵容忍系统的现状与缺点

入侵容忍体系构造中一个重要的组成部分就是入侵容忍触发器的构造,即在容忍入侵的条件下如何判断一个入侵行为的发生。目前主要有两种实现方式:(1)将IDS进行改进作为容忍入侵的触发检测机制<sup>[2-4]</sup>,但这种方法对于“自我”缺乏更新性,除此之外,“记忆细胞”的非死亡性,可能会导致触发器的误报率显著增加,并对新的入侵类型容易忽视;(2)以表决的结果作为容忍入侵触发的机制<sup>[5]</sup>,其核心思想是利用冗余和多样性技术进行表决进而判断系统是否受到了攻击,这种方法的最大缺点是会占用大量的空闲资源,而且也会导致很高的误警率。

为此本文将突变原理加入到免疫记忆中提出了一种基于生物免疫突变的入侵容忍触发模型(A New Instrusion Tolerance Trigger Model Based on Biological Immune and Catastrophe Theory, BICTTM)。主要思想是通过突变原理的势函数和临界点来判别人侵,当有入侵行为发生时由于存在容忍机制故

不必马上报警进行处理,而随着攻击行为的深入如果达到了一个临界值或者达到某一条件后再进行报警处理,以此来提升系统抵抗入侵的能力,达到入侵容忍的效果。

## 2 BICTTM 模型

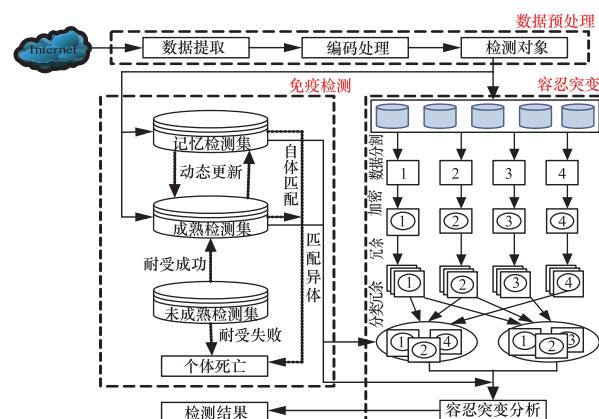


图1 生物免疫突变的入侵容忍触发模型框架

### 2.1 模型框架

自框架图可以看到 BICTTM 模型由三个部分组成:

(1) 数据预处理。此过程负责收集需要检测对象的相关信息,进行离散化编码,按照一定的编码

存放规则生成检测对象。

(2) 免疫检测。其主要是模拟人工免疫系统的否定选择、自体耐受、克隆选择以及细胞记忆等过程,由记忆检测集和成熟检测集对需要检测的对象进行分类,以发现异常数据。

(3) 容忍突变分析。根据容忍机制和容忍强度对入侵的攻击行为进行分析并做出预警或相应的处置措施。

## 2.2 模型定义

在网络系统中,进程调用各种资源具有相对稳定的行为特征,因此本模型中,以进程对资源的占有、消耗、放弃等行为特征来进行相应的信息处理。

**定义 1** 检测集  $V$  是一个四元组:

$$V_s = \{ <(P_{ID}, P_{path}, P_{file}, R_V), R> | ; \\ P_{ID} \in P, P_{ID, path, file, v}, R \in D \} \quad (1)$$

检测集  $V$  用来表示服务器系统运行中以及进程对系统资源的使用情况,式(1)中: $D = \{0, 1\}^L$ ,  $|V| = L$ ,  $L$  为自然常数。和生物的免疫系统相类似,本模型的检测对象编码格式也由不变区和可变区组成,在不变区中,  $P_{ID}$  表示进程编号,  $P_{path}$  表示进程的路径,  $P_{file}$  表示文件的使用程度,  $R_V$  表示进程的资源类型;可变区域由  $R$  表示进程资源使用序列构成,其序列长度是可变的。

$S \in \{normal, unNormal\}$  表示检测集的类型,其中  $V_{normal}$  表示正常进程及进程的资源使用情况(自体);  $V_{unNormal}$  表示异常进程及进程资源的使用情况,且满足如下条件:  $V_{normal} \cup V_{unNormal} = V$ ,  $V_{normal} \cap V_{unNormal} = \phi$ 。

**定义 2** 检测因子集  $E_s$  为一个三元组:

$$E_s = \{ (B_{sequence}, K_{time}, K_n) | K_{time} \in N, K_n \in N \} \quad (2)$$

式(2)中: $B_{sequence}$  为检测因子的检测序列(抗体),其编码格式与  $V$  相同,由其完成对检测对象的相似匹配; $K_{time}$  为检测因子的生存时间; $K_n$  为匹配计数器。

$S \in \{ripe, memory\}$  分别表示成熟因子检测集  $E_{ripe}$  和记忆检测集  $E_{memory}$ ,成熟因子是指不成熟的因子与  $V_s$  匹配反映发生后,不对  $V_s$  产生识别的,且在检测  $V_s$  时未与  $V_s$  中的检测对象相匹配的检测因子集;记忆检测集  $E_{memory}$  是指成熟检测因子在检测  $V_s$

时,与检测对象发生匹配且  $K_n$  超过阈值的检测因子集。定义如下:

$$\begin{aligned} E_{ripe} = \{x | x \in E_s, \forall y \in V_s, &<x, B_{sequence}, \\ &y> \notin F \wedge x, K_n < \eta\} \end{aligned} \quad (3)$$

$$\begin{aligned} E_{memory} = \{x | x \in E_s, \forall y \in V_s, &<x, B_{sequence}, \\ &y> \notin F \wedge x, K_n \geq \eta\} \end{aligned} \quad (4)$$

它们满足关系:

$$E_{ripe} \cup E_{memory} = E_s, E_{ripe} \cap E_{memory} = \emptyset, F \text{ 为匹配集合}^{[4]} \\ F = \{ <x, y> | x, y \in D, f(x, y) = 1 \} \quad (5)$$

**定义 3** 未成熟检测因子集  $E_{unRipe}$  为一个二元组:

$$E_{unRipe} = \{ (B_{sequence}, K_{time}) | K_{time} \in N \} \quad (6)$$

其表示检测器按照系统的检测因子  $E$  的编码格式随即生成的且必须经过与  $V_s$  发生识别反映,用于生成  $E_{ripe}$  的检测因子集。

## 2.3 更新与突变检测算法

### 2.3.1 检测对象集的自动更新

对象检测集必须实时地检测网络系统中的进程及其占用资源的动态情况,本模型采用和文献 [4] 相类似的以时间为轴的动态更新方式。

$V_s$  的更新过程如下:

$$\left\{ \begin{array}{l} V_s(t+1) = V_s(t) - V'_s(t) + \Delta V_s(t) \\ V'_s(t) = \{x | x \in V_s(t), \exists y \in E_s(t), (x, y, B_{sequence}) \in F\} \\ \Delta V_s(t) = \{x | x \in V_s(t), \exists y \in E_s(t), (x, y, B_{sequence}) \notin F\} \end{array} \right. \quad (7)$$

式(7)中: $\Delta V_s(t)$  表示在  $t$  时段内新增的检测对象集,  $V'_s(t)$  表示  $t$  时段内  $V_s$  发生了改变且不再属于  $V_s$  的检测对象。

### 2.3.2 未成熟检测因子的生成与更新

未成熟检测因子是由检测器自动随机生成的,其中一部分会死亡(如果与  $V_s$  相匹配),另一部分会在一定的时间内成长为成熟检测因子。 $E_{unRipe}$  的更新过程如下:

$$\begin{aligned} E_{unRipe}(t+1) = & E'_{unRipe}(t) - E_{ripe'}(t) + \Delta E_{unRipe}(t) \\ E'_{unRipe}(t) = & \{y | \exists x \in E_{ripe}(t), y, B_{sequence} = \\ & x, B_{sequence}, y, K_{time} = x, K_{time} + 1\}; \\ E_{ripe'}(t) = & \{x | x \in E_{ripe}(t), x, K_{time} > \varepsilon\} \end{aligned} \quad (8)$$

式(8)中: $E'_{unRipe}(t)$  表示当前时刻剩余的未成熟检

测因子集;  $E_{ripe'}(t)$  表示在  $t$  时段内变为成熟的检测因子;  $\Delta E_{unRipe}(t)$  表示在本时段随机生成的未成熟检测因子。

### 2.3.3 成熟检测因子的动态更新

成熟检测因子的作用是对检测对象  $V_s$  进行分类识别,当判别某一个对象为  $V_{unNormal}$  时即进入到容忍突变检测过程中。

$E_{ripe}(t) = \{x_1, x_2, \dots, x_{|E_{ripe}|}\}$  表示  $t$  时刻的成熟检测因子集,则其动态的更新过程如下:

$$\begin{aligned} E_{ripe}(t+1) = & E'_{ripe}(t) + \Delta E_{ripe}(t) + E_{r-m}(t) + \\ & E_{m-r}(t) - E_{ripe_m}(t) - E_{ripe_\gamma}(t) \end{aligned} \quad (9)$$

式(9)中: $E'_{ripe}(t)$  表示在  $t$  时刻内仍然保持着的成熟检测因子,同时更新其生存周期; $\Delta E_{ripe}(t)$  表示在  $t$  时段内由未成熟检测因子检测而演变来的成熟检测因子; $E_{m-r}(t)$  表示由记忆检测因子降级而来的成熟检测因子; $E_{ripe_m}(t)$  表示在  $t$  时段内转化为记忆检测因子的成熟检测因子集; $E_{r-m}(t)$  表示发生了变化的新的成熟检测因子集; $E_{ripe_\gamma}(t)$  表示在  $t$  时段内由于生存周期  $\gamma$  内未与检测对象发生匹配作用的成熟检测因子集。

### 2.3.4 记忆检测因子的更新

记忆检测因子的作用就是当  $V_{unNormal}$  中的对象在此发现时,其可以快速的识别,而且有更长的生命周期,以此来提升检测效率。

假设  $E_{memory}(t) = \{x_1, x_2, \dots, x_{|E_{memory}|}\}$  为  $t$  时刻的记忆检测因子集合,其更新方式为:

$$E_{memory}(t) = E'_{memory}(t) + \Delta E_{memory}(t) - E_{m-e}(t) - E_{m-r}(t) \quad (10)$$

各元素意义类同式(9)。

### 2.3.5 容忍突变检测分析

在本模型中考虑系统已经检测到流量攻击,并处于抵制状态的情况,此时需要判断系统是否仍能够正常运作或者确定报警。

根据突变理论需要依据容忍系统变量间的关系函数来确定突变点,在进程对资源的使用过程中,数据的传输可以看作是一系列数据包的有序传送,动能函数近似等于所有数据包的动能,所以本

模型利用动能函数来反应系统的势能,全局势能函数定义如式(11)。

$$G = \sum_i \frac{1}{2} k B_i(t) u_i^2(t) \quad (11)$$

其中  $k$  为标准单位数据包大小系数,则容忍势函数可表示为:

$$G = \sum_i \frac{1}{2} k B_i(t) u_{i0}^2 \left( 1 + \alpha_i \left( \frac{B_i(t)}{C_i(t)} \right)^{\beta_i} \right)^{-2} \quad (12)$$

同时局部容忍势函数为:

$$F_i(B_i, C_i) = \frac{1}{2} k B_i(t) u_{i0}^2 \left( 1 + \alpha_i \left( \frac{B_i(t)}{C_i(t)} \right)^{\beta_i} \right)^{-2} \quad (13)$$

突变临界点集可由  $\frac{\partial F_i(B_i, C_i)}{\partial B_i} = 0$  求得。即:

$$\frac{\partial F_i(B_i, C_i)}{\partial B_i} = 2\alpha_i \beta_i \left( \frac{B_i}{C_i} \right)^{\beta_i} - \alpha_i \left( \frac{B_i}{C_i} \right)^{\beta_i} - 1 = 0 \quad (14)$$

也即:  $\frac{B_i}{C_i} = \left( \frac{1}{2\alpha_i \beta_i - \alpha_i} \right)_i^{\beta_i-1}$  如果式(14)大于 0, 则说明相应的突变临界点是稳定的,若式(14)等于 0, 则为容忍势函数关于数据流量的拐点。

当稳定点与分叉集点重合时,容忍势函数的曲率最大,即系统出现突变后,只要限制数据流量小于或等于稳定点数据流量,就能保证系统的容忍性并提供稳定服务,否则证明系统存在着严重的入侵问题,需要报警或者采取相应的措施。

## 3 实例分析

为了能够真实地模拟和验证本模型的正确性和效率,本次试验和文献[4]中的处理方式一致,使用常用的标准测试数据集 UCI(University of California, Irvine)数据集中的分类验证数据作为样本集。(下载地址 <http://kdd.ics.uci.edu/summary.data.type.html>)。

在实验的过程中,本模型首先按照类别进行分组,形成对象检测集  $V_s$ ,对象集中共有 16 个属性值,其包括:进程编号,进程路径、进程文件占用量、访问时间、访问的 IP 地址、CPU 占用率、网络流量、网络吞吐量、I/O 占用的时间等,其中进程编号为固

定的,其余会随之而改变,在进行实验时采用相应的二进制编码,进行相应的处理,得出模型的异常识别率和正常识别率如图 2 所示。

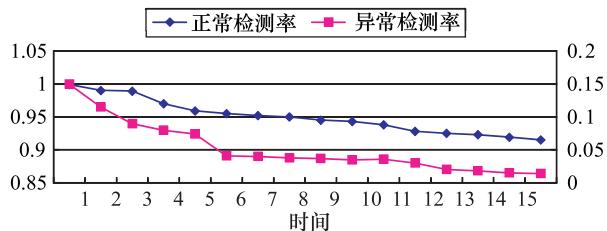


图 2 实验检测率结果图

实例证明,本模型能够对正常的进程进行检测,而且由于对象检测集  $V_s$  的动态更新,因此具有较高的检测效率,正常检测率在 0.9 以上,这较以往的模型有很大的优势。例如传统的 Dynamics 算法

的正常检测率最多达到 0.86 左右。但由于采用的容忍突变机制使得系统的异常检测率有所增加,需要在以后的研究中进一步改进。

## 参 考 文 献

- 1 张鸿志,张玉清. 网络可生存性研究进展. 计算机工程, 2005; 31(20):350—355
- 2 刘丽,刘传忠,王宏,等. 基于诱骗机制的网络容侵模型的设计与实现. 计算机工程与设计, 2006; 27(8):1435—1438
- 3 张萍,王健忠. 基于教育网格的免疫安全考试系统. 计算机应用, 2006; 26(2):349—351
- 4 刘冠. 免疫原理在入侵容忍系统中的应用研究. 西安建筑科技大学硕士论文, 2007
- 5 郭健. 基于突变理论的复杂系统的脆弱性研究. 哈尔滨:哈尔滨工程大学, 2002; 12:12—13

## A Intrusion Tolerance Trigger Model Based on Biological Immune and Catastrophe Theory

LEI Li-ping

(Weinan Municipal Party Committee Party School, Weinan 714000, P. R. China)

[Abstract] The intrusion detection system is one of the key technologies for network safety. Smart technology is one important branch, and the detection technology intellectualized is an important development. But the present systems are not perfect. Based on the biological memory principles, Biological Immune and Catastrophe Theory etc are combined into the present intrusion detection system frame, and propose a new intrusion tolerance trigger model based on biological immune and catastrophe theory, in order to make the invade detection system more intelligent.

[Key words] memory principles      intrusion detection      tolerate mutation      biological immune principles