

基于免疫重构的阴性选择算法^{*})

罗一丹 蔡自兴 王 勇 江中央

(中南大学信息科学与工程学院 长沙 410083)

摘 要 针对阴性选择算法在解决实际问题中,易误判及自修复能力差的弱点,该文基于生物免疫系统内部学习优化机制以及工业领域中的可重构系统,提出了一种基于免疫重构的阴性选择算法。新算法将可重构系统的思想融入到阴性选择算法中,提出了重构串、重构模型与重构操作的概念与实现方法,以保证系统发生意外的时候能够及时恢复、重组。将算法应用于一个 Web 系统进行仿真实验,结果表明该算法是有效的。

关键词 阴性选择算法,重构串,重构操作,重构模型

Negative Selection Algorithm Based on Immune Reconfiguration

LUO Yi-Dan CAI Zi-Xing WANG Yong JIANG Zhong-Yang

(College of Information Science and Engineering, Central South University, Changsha 410083)

Abstract Considering the weakness of self-repairing and misjudgment of negative selection of immune algorithm, a new negative selection algorithm based on immune reconfiguration is proposed. Based on the interior study mechanism of biological immune system and reconfigurable system of industrial field, the new algorithm proposes the conceptions and application methods of reconfigurable string, reconfigurable operator and reconfigurable model and integrates the idea of reconfigurable system into negative selection algorithm. In the case of accidents the system can be recovered and reformed in time. A Web system is utilized to test this method and the simulation results suggest that this algorithm has good performance.

Keywords Negative selection algorithm, Reconfigurable string, Reconfigurable operator, Reconfigurable mode

生物免疫系统是一个高度并行、分布、自适应和自组织的系统。基于生物免疫系统的人工免疫模型和算法研究已经成为计算智能领域中继人工神经网络、进化计算之后又一个新的研究热点。而人工免疫系统(Artificial Immune System, AIS)是一类基于生物免疫系统(Biological Immune System, BIS)功能、原理、基本特征以及相关理论免疫学说而建立的用于解决各种复杂问题的计算系统^[1,2]。它实现了一种受生物免疫系统启发,通过学习外界物质的自然防御机理的学习技术。人工免疫系统算法包括基于免疫网络学说的人工免疫网络模型,基于免疫特异性的阴性选择算法^[3](Negative Selection Algorithm, NSA),克隆算法^[4](Clone Algorithm)和免疫进化算法^[5](Immune Evolutionary Algorithm)。下面我们在分析免疫算法中的阴性选择算法的基础上,提出一种基于免疫重构的阴性选择算法(Negative Selection Based On Immune Reconfiguration, NSBOIR)。

1 阴性选择算法

生物免疫系统通过组织淋巴细胞来识别“自体”和“非体”抗原,消灭并清除异物。在生物免疫系统中,淋巴细胞分为 T 淋巴细胞和 B 淋巴细胞。在抵御病菌入侵时, T 淋巴细胞执行特异性免疫反应和调节功能。B 淋巴细胞则在其表面产生抗体,探测对应的抗原并记录下该抗原的结构和特征,执行识别、记忆、学习等功能。美国 New Mexico 大学的 Forrest 教授模拟淋巴细胞的产生过程,提出阴性选择算法^[3,6,7]。

阴性选择算法采用基于免疫系统中的自体 and 异物识别的原理来进行变化监测,先随机产生监测器,然后删除与自身对

抗的并保留能检测异体的细胞。如图 1 所示,算法过程^[3]如下所述:

步骤 1 首先生成长为 L 的预检测器,然后与自我串集合按照匹配规则进行匹配,若匹配成功,则删除;否则,放入成熟检测器集合;

步骤 2 重复以上过程,直至生成预定数量的成熟检测器;

步骤 3 成熟检测器用于匹配待检测串,若匹配,则表明发生了异常变化。

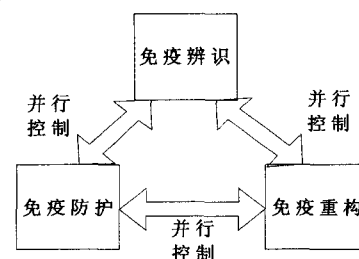


图 1 阴性选择算法流程图

在实际应用中,阴性选择算法在检测时存在着误判、误删的问题,将异常文件误认为是正常文件而不做相应的处理或将正常的文件认为是异常而删除,使得整个系统发生异常甚至致使系统崩溃。针对这个问题,本文利用重构的优点——可对已经发生的错误进行一个返回并对文件进行重组,使文件回到一个正常的组态,将重构融入到阴性选择算法中产生一种新的算法——基于免疫重构的阴性选择算法(NS-

^{*}国家自然科学基金资助项目(60402021,60234030),国家基础研究项目(A1420060159)及湖南省院士基金(05JJY3035)。罗一丹 硕士研究生,主要研究方向:免疫优化算法,免疫算法;蔡自兴 教授,博士生导师,主要研究领域:人工智能,智能控制,智能机器人和自动控制等;王 勇 博士,讲师,研究方向:进化计算等;江中央 硕士研究生,主要研究方向:免疫优化算法,免疫控制。

BOIR)。

2 基于免疫重构的阴性选择算法 (Negative Selection Based On Immune Reconfiguration)

2.1 基本概念

2.1.1 重构的定义

重构在工业生产领域已有较频繁的涉及如:制造业^[8]。

重构指系统从一种构形向另一种构形的转换,是系统生存和发展的基本手段。

2.1.2 重构的模型

传统的阴性选择算法,在检测时可能会出现误判、误删等问题。而重构能够对已经发生的错误进行一个返回,使系统恢复到正常组态。它首先对计算机的异常进行辨识以判断是文件任务还是意外事件,若是文件任务,则不做其它处理;否则,对意外进行实时处理,并对文件进行重组使之恢复到一个正常的组态。重构的模型如图 2 由以下几部分组成:

(1)免疫辨识。监控和分析文件数据,并将检测到的文件的任务和意外事件信息正确分发给重构和免疫防护进行处理。

(2)免疫防护。对意外及突发事件进行实时处理,对基于文件的异常进行处理。

(3)免疫重构。根据外界需求重构,当系统发生不可逆转的错误时,对文件进行重组,形成逻辑构型新组态。

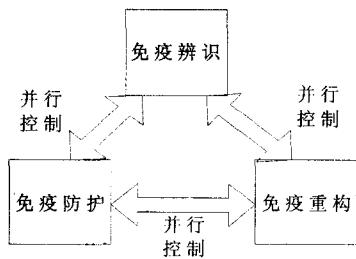


图 2 重构的模型

2.2 算法的基本思想

本文提出的基于免疫重构的阴性选择算法,不仅采用了基于免疫系统中的自体和异体识别的原理来进行变化监测,并且加入了工程应用中基于构形转换的重构思想。算法的基本思想为:首先随机产生监测器即随机串的集合,其次删除与自身对抗的,并保留能检测异体的细胞,生成检测器集合;然后生成重构串,与待检测串进行匹配,判断是文件任务还是意外事件。

2.3 NSBOI 重构模型描述

2.3.1 重构串 (Reconfigurable string)

定义 1(重构串) 假设发生一场意外事件,用户认为有重构的需要,可以执行一个操作来促发重构,此操作的编码即为重构串。

2.3.2 重构操作 (Reconfigurable operator)

定义 2(重构操作) 假设所有的待检测串、自体串、重构串集合的串都是一种编码方法,存储每一个待检测串并记录其来源(文件)及最后修改时间,文件是有备份的。重构操作包括:

(1)在待检测串集合中离重构串的操作时间最短的串称为异常串,并将其放入检测器中。这样,检测器就保留了异体,对此异体免疫。

(2)将异常串的来源文件删除。

(3)将备份的文件,覆盖现有的文件。

2.3.3 NSBOI 重构模型 (NSBOI Reconfigurable model)

定义 3(NSBOI 重构模型) NSBOI 模型包括:重构串、重构操作等。参照图 2 分为三部分:

(1)由重构串来辨识是文件的任务还是意外事件。

(2)免疫防护。此系统将对程序编码形成的待检测串与检测器进行匹配,如果两者之间的亲合度高于某一阈值,则认为该抗原是异体,发出报警信息。

(3)进行重构操作,来完成文件的重组。

2.4 算法的步骤

如图 3 所示,基于免疫重构的阴性选择算法的步骤如下:

Step 1 首先生成长为 L 的预检测器,然后与自我串集合按照匹配规则进行匹配,若匹配成功,则删除;否则,放入成熟检测器集合。

Step 2 重复以上过程,直至生成预定数量的成熟检测器。

Step 3 生成长度为 L 的重构串集合。

Step 4 将检测串与重构串匹配,若匹配成功,则转 Step 5;否则,与检测器集合匹配,若匹配,则表明发生了异常变化。

Step 5 按照 NSBOIR 重构模型,对检测器集合,进行重构操作,并对文件进行重组、恢复。

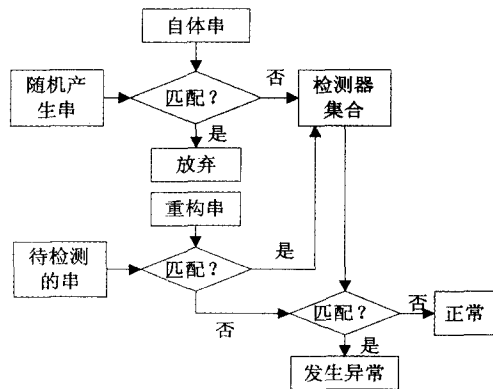


图 3 NSBOI 算法流程图

3 基于免疫重构的阴性选择算法在 Web 系统中的应用

对于电脑空间,每个组件在其四维时空坐标系^[9]里也应该有唯一确定的时空坐标点,也就是说每个组件都由其时空属性唯一确定^[9]。在 Web 系统中,各文件都有其正常的状态。但是当意外事件发生时,阴性选择算法在检测的时候可能会误判,将正常的文件认为是病毒而删除或将异常文件误认为是正常文件而不做报警。

3.1 Web 系统的正常状态^[10]

假设 Web 系统 S 由 n_d 个 Web 文件夹和 n_f 个 HTML 网页文件组成,该系统是静态的,那么该系统的文件集合表示为:

$$\{c_{i,j} | c_{i,j} \in S, i=1, \dots, n_d, j=1, \dots, n_i, \sum_{i=1}^{n_d} n_i = n_f\}$$

这里, $c_{i,j}$ 表示系统 S 的第 i 文件夹中的第 j 个文件, n_i 表示系统 S 的第 i 文件夹中所有文件的总数,这些文件夹和文件称为该系统的组件。

因而,所述 Web 系统的正常状态表示为其所有组件的正常状态。受自然免疫系统的正常状态启发,该 Web 系统的正常状态表示为其正常组件的时空属性。该系统中每个组件 $c_{i,j}$ 的空间属性 p_i 表示为该组件的绝对路径名,该组件的时间属性表示 t_i 为该组件的最后修改时间。那么所有组件时

空属性的集合可表示为:

$$(D, T) = \{(p_i, t_i) | p_i \in c_{i,j}, t_i \in c_{i,j}, i=1, \dots, n_d + n_f\}$$

定理 1 假设 Web 系统 S 由 n_f 个 HTML 网页文件和 n_d 个文件夹组成, 该系统中每个组件 c_i 的空间属性 p_i 表示为该组件的绝对路径名, 该组件的时间属性 t_i 表示为该组件的最后修改时间。那么, 正常系统 S 中所有正常文件的绝对路径名和最后修改时间唯一确定该系统的正常状态。

由此, 我们可以看出, 一个 Web 系统由多个文件和文件夹组成, 这些文件与文件夹称为该系统的组件, 而每个组件 $c_{i,j}$ 都有一个唯一与之对应的时空属性 (p_i, t_i) 即组件与时空属性形成一个一一映射。

3.2 算法的应用

3.2.1 文件的定义

由上述说明我们可以得知每个组件都有唯一的 $(c_{i,j}, p_{i,j}, t_i)$ 。因此, 本文选择了文件操作的 3 个参数: 组件名, 组件的绝对路径名, 组件的最后修改时间, 并对其编码, 以方便操作。

3.2.2 随机串 A_k 的编码

采用二进制编码, 并加上文件的定义来组成。如: $A_k = ((a_{k,1}, a_{k,2}, \dots, a_{k,N}), c_{k,i,j}, p_{k,i,j}, t_k), k=1, 2, \dots, n$ 。N 表示串的长度; n 表示串的规模; $c_{k,i,j}$ 表示 A_k 串是属于第 i 文件夹中的第 j 个文件; $p_{k,i,j}$ 表示 A_k 串是第 i 文件夹中的第 j 个文件的绝对路径; t_k 表示 A_k 串的最后修改时间。可以看出, 随机串的规模, 主要是由检测器的规模决定, 一旦生成了一定规模的检测器, 就可停止生成随机串, 所以本文将检测器规模定在 10^2 。

3.2.3 匹配规则

首先 A_k 中的第一项 $(a_{k,1}, a_{k,2}, \dots, a_{k,N})$ 的匹配规则, 按阈值 T 来控制, 如果匹配度大于 T 则认为是匹配, 否则, 认为不匹配。

随机产生串与自体串、待检测串与重构串、抗原与检测器的匹配原则统一为 A_k 中的前三项 $(a_{k,1}, a_{k,2}, \dots, a_{k,N}), c_{k,i,j}, p_{k,i,j}$ 之间的匹配。首先匹配 $c_{k,i,j}$, 再匹配 $p_{k,i,j}$, 最后按 A_k 中的第一项匹配规则匹配 $(a_{k,1}, a_{k,2}, \dots, a_{k,N})$ 。

3.2.4 重构操作

假设串 $A_k = ((a_{k,1}, a_{k,2}, \dots, a_{k,N}), c_{k,i,j}, p_{k,i,j}, t_k), k=1, 2, \dots, n, (A_k, i)$ 就表示 A_k 集合中的第 i 项。设重构发生的时间为 t, 对所有待检测串中 t_k , 假若 $k=h$ 时, 可以使得 $\min |t - t_k|, k=(1, 2, \dots, n)$ 。用同样的方法找到 t_l 使得 $k=1$ 时 $\min (|t_h - t_k|)$ 。

将 $(A_h, 1) = (a_{h,1}, a_{h,2}, \dots, a_{h,N})$ 加入到检测器集合, 删除所有 $(A_k, 4) = t_h$ 的文件、文件夹以及对系统的任何操作。所有组件全部替换成 $t=t_l$ 时的组件。

3.2.5 实验

本文基于 Windows 操作系统平台, 使用 Visual C++ 6.0 开发工具和面向对象的方法设计开发了程序, 如图 4 所示, 上面一些按钮是用来控制 NSBOI 与 NSA 的操作按钮, 分别用来启动阴性选择算法程序, 启动基于免疫重构的阴性选择算法程序, 打印实验结果和保存日志。

当传统的阴性选择算法产生误判时, 启动基于免疫重构的阴性选择算法, 弹出一个对话框询问是否需要重构, 选择 yes, 结果如图 4 所示。

100 次反复实验结果表明, 当所述 Web 系统无病毒感染时, 所述基于免疫重构的阴性选择算法程序将所有的文件和文件夹检测为自体, 并不会发现异体, 不存在将自体检测为异体的情况; 当所述 Web 系统受病毒攻击时, 阴性选择算法能

将意外及其感染的文件检测为异体的概率为 60%, 误判后, 运用基于免疫重构的阴性选择算法消除异体、修复受损的 Web 系统, 使系统恢复到上一次正常的情况的概率为 90%。上述实验结果说明了基于免疫重构的阴性选择算法同传统的阴性选择算法相比, 较大地提高了消除意外的能力。该技术能够实现受损 Web 系统的修复, 为 Web 系统的网络安全和用户自主维护提供了有效的、强大的技术支持。

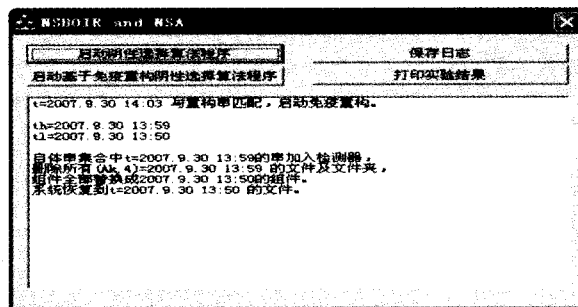


图 4 NSBOI and NSA 程序图

结论 本文提出了一种基于生物免疫系统内部学习优化机制以及可重构系统的基于免疫重构的阴性选择算法 (NSBOIR), 同时对该算法产生机理、算法中各个关键创新进行了讨论。该算法随机产生检测器即随机串的集合; 删除与自身对抗的并保留能检测异体的细胞, 生成检测器集合; 然后生成重构串, 与待检测串匹配, 判断是文件任务还是意外事件。本文还提出一系列新的概念: 重构串、重构模型、重构操作, 并且对一个 Web 系统进行实验测试, 结果表明该算法是有效的, 它不仅可以对异体进行报警, 更重要的是当系统发生意外时, 可以立即重新调整其组件, 恢复系统并对此意外免疫。

参考文献

- Hofmeyr S, Forrest S. Immunity by design: An artificial immune system. In: Wolfgang B, Jason M D, et al, eds. Proc. of the Genetic and Evolutionary Computation Conf. San Francisco: Morgan Kaufman Publishers, 1999. 1289~1296
- De Castro L N, Timmis J. Artificial Immune Systems: A New Computational Intelligence Approach. Heidelberg: Springer-Verlag, 2002
- Forrest S, Perelson A S, Allen L, et al. Self-nonsel self discrimination in a computer. In: Proceedings of the IEEE Symposium on Research in Security and Privacy. Los Alamitos: IEEE Computer Society Press, Los Alamitos, CA, 1994. 202~212
- De Castro L N, von Zuben F J. The clonal selection algorithm with engineering applications. In: Whitley L D, Goldberg D E, et al, eds. Proc. of the GECCO 2000. San Francisco: Morgan Kaufman Publishers, 2000. 36~37
- Liu Ruo-Chen, Du Hai-Feng, et al. Clonal Strategy Algorithm Based on the immune memory. Journal of Computer Science and Technology, 2005, 20(5): 728~734
- Forrest S, Hofmeyr S A, Somayaji A, et al. A sense of self for UNIX processes. In: Proceedings of the 1996 IEEE Symposium on Security and Privacy. Los Alamitos, CA, 1996. 120~128
- D'haeseleer P, Forrest S. An immunological approach to change detection algorithms analysis and implications. In: Proceedings of the IEEE Symposium on Security and Privacy. LOS Alamitos, CA, 1996. 110~119
- 梁福军, 宁汝新. 可重构制造系统理论研究. 机械工程学报, 2003, 39(6): 36~43
- Einstein A. Relativity: the Special and General Theory. New York: Three Rivers Press, 1920
- 龚涛. 免疫计算的建模与鲁棒性分析研究: [博士学位论文]. 湖南: 中南大学, 2007