第34卷 第23期 计 算 机 工 程 2008年12月
Vol.34 No.23 Computer Engineering December 2008
·安全技术· 文章编号：1000—3428(2008)23—0164—03 文献标识码：A 中图分类号：TP309

# 基于免疫网络的信息安全风险评估模型

黄 欢，庄 毅，许 斌
(南京航空航天大学信息科学与技术学院，南京 210016)

**摘 要：**风险评估是评价网络信息系统安全的有效措施之一。该文基于免疫网络可动态实时诊断的特性，提出一种新的信息安全风险评估模型，给出模型中各项指标的定量计算方法，以评估整个信息系统的风险值。该模型能够综合考虑评估要素的相互关联，针对风险动态更新，进行实时监控。实验验证了其评估信息系统安全状态的有效性。

**关键词：**风险评估；免疫网络；风险量化

# Risk Assessment Model for Information Security Based on Immune Network

HUANG Huan, ZHUANG Yi, XU Bin
(Institute of Information Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016)

**【Abstract】** Risk assessment is an effective approach to evaluate security state of information systems. A model of information security risk assessment system based on immune network both with a quantitative evaluation theory are presented. This model can reflect the relationship of the evaluation factors. It has the features of self-learning, self-update and real-time detecting. Experiment shows the model is effective to assess the risk of information systems.

**【Key words】** risk assessment; immune network; risk quantification

## 1 概述

信息安全风险评估是一种系统性研究，通过考察信息系统安全所涉及的各个方面，评估信息系统的安全性能和风险状态，重点评估系统核心信息资源的安全质量状态等级值。目前，Summers, Carroll, Pfleeger等已提出了各自的风险评估方法，一些安全公司和科研团体也推出了诸如CRAMM, COBRA, OCTAVE等风险评估模型和工具[1]。这些方法和工具大多采用静态的评估方法，着眼于管理层面对信息系统作宏观的评价，无法满足细化、定量风险分析的要求。

本文提出一个定量的网络安全风险评估系统模型。该模型的特点表现在以下方面：(1)其理论基础建立在一个公认的风险评估方程[2]上。(2)对威胁事件、资产价值、脆弱性程度进行量化计算，给出了定量评估系统风险的方法。(3)基于免疫网络模型，可以实时监测网络面临的攻击，具备自学习、自增长的能力，能够应对不断变化的系统和网络环境。

## 2 风险评估系统模型

由国家信息中心信息安全研究与服务中心、国家保密技术研究所、中科院信息安全国家重点实验室等多家机构合作起草的《信息安全风险评估指南》给出了风险评估的流程和总体框架，提出一种风险计算的概念方法[3]：

$$\text{风险值} = R(A,T,V) = R(L(T,V),F(Ia,Va)) \quad (1)$$

其中，$R$ 表示安全风险计算函数；$A$ 表示资产；$T$ 表示威胁；$V$ 表示脆弱性；$Ia$ 表示安全事件所作用的资产重要程度；$Va$ 表示脆弱性严重程度；$L$ 表示威胁利用资产的脆弱性导致安全事件发生的可能性；$F$ 表示安全事件发生后的损失。

式(1)给出了风险评估中所有的指标，但还停留在抽象的理论层面，对安全事件发生可能性的计算、损失的衡量等问题没有深入地研究，在实际应用中缺乏可操作性。因此，本文对该风险计算方法进行了细化，并设计了信息安全风险评估计算模型，见图1。

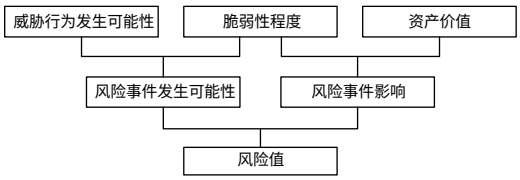


**图1 信息安全风险评估计算模型**

在图1中，风险事件发生可能性受威胁行为发生可能性、脆弱点的脆弱程度2个因素影响，风险影响值则与资产价值和脆弱点的脆弱程度密切相关。

根据上述风险计算模型，设计一个基于免疫网络的信息安全风险评估系统，主要分为4大功能模块，如图2所示。其中，资产检测模块负责对每个节点的资源进行评估，确定它们对于整个信息系统的重要性和安全需求等级；威胁行为检测模块负责对节点当前所面临的外在攻击进行检测，记录威胁发生所利用的脆弱点和遭受威胁的资产，利用量化模型计算威胁值；脆弱性检测模块负责评估信息系统中各个节点的漏洞，通过量化模型转换成节点的脆弱性指数。这3个模块的核心部分是各自的检测器集合，作为信息系统的抗体集，

**基金项目：**航空科学基金资助项目(05F2037, 04c52009)；国家部委预研基金资助项目

**作者简介：**黄 欢(1983－)，女，硕士研究生，主研方向：网络安全；庄 毅，教授；许 斌，博士研究生

**收稿日期：**2008-06-17 **E-mail：**xoxihuan2001@sohu.com

它们共同构成风险评估免疫网络，实现对信息系统风险的实时、动态免疫功能。风险计算模块则根据上述 3 个模块的输出结果，用量化模型计算出整个系统的安全风险值。

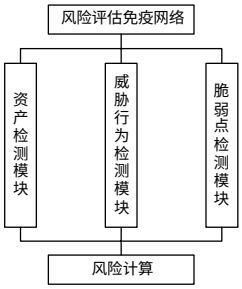


**图 2　信息安全风险评估系统框架**

模型采用客户/服务器(C/S)模式。服务器端运行风险计算程序，负责对目标系统进行资产评估和配置管理，根据客户端返回的检测结果对目标主机或网络进行安全风险的评估。客户端运行风险评估免疫网络程序，主要负责威胁检测、资产检测和脆弱性检测。服务器端和客户端之间采用证书认证机制进行安全认证。

## 3　基于免疫网络的风险评估关键技术

免疫网络的原理主要建立在自身识别上。免疫系统淋巴细胞上分布的特异性抗原受体可变区(V)组成内网络，通过免疫细胞相互识别 V 区上的抗原决定簇来实现免疫系统的功能。免疫网络中各个细胞克隆并非处于一种独立状态，而是通过自我识别、相互刺激和相互制约构成一个动态平衡的网络结构。

目前的免疫网络学说主要以文献[4]提出的资源限制的人工免疫网络学习算法 RAIN 和文献[5]提出的免疫网络模型算法 aiNet 为代表，这 2 种算法都运用了免疫网络自我平衡和调节的机理，但前者采用了实数编码，数据使用前需要标准化，由此产生的影响还未知；而后者存在的不足是参数较多，计算成本高，网络对为了控制规模而预先设定的规模抑制阈值较为敏感。还有一些基于免疫网络机理提出的免疫进化算法则大多采用海明距离或信息熵的方法求出抗体浓度，再根据浓度进行选择操作。这类算法收敛速度较慢，且没有完全体现免疫网络的调节机理，对算法系统的数学理论分析也比较少。本文从免疫网络的基本理论出发，在继承了免疫应答、遗传变异、克隆、记忆、动态平衡等特性的基础上，设计了适用于信息系统风险评估的免疫网络模型。

### 3.1　免疫网络的构建

威胁行为必须利用系统中一个或多个脆弱点达到其危害的目的，威胁行为的危害必然是针对系统中的一个或多个资产。根据信息系统中威胁行为发生的逻辑关系，可以构建免疫网络。相互之间有关联的检测器在网络中是联通的。没有逻辑关系的检测器不联通。设资产检测器集合 $Da\{Da1, Da2, \cdots, Dan\}$、威胁行为检测器集合 $Dt\{Dt1, Dt2, \cdots, Dtm\}$、脆弱点检测器集合 $Dv\{Dv1, Dv2, \cdots, Dvk\}$及 $m,n,k$ 均为自然数。一个信息安全免疫网络的示例如图 3 所示。

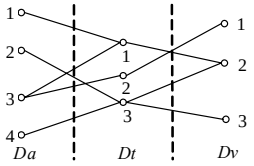


**图 3　免疫网络的示例**

在图 3 中， $Dt3$ 针对的威胁行为通过利用 $Dv2$ 和 $Dv3$ 检测到的脆弱点来危害 $Da2$ 和 $Da4$ 所检测的资产。

### 3.2　检测器生成

根据该系统中 3 种检测器的不同功能，将其分为动态检测器和静态检测器。威胁行为检测器为动态检测器，在免疫网络中处于核心地位，检测信息系统所遭受威胁行为的种类、强度、所用脆弱点、受危害资产等指标，将结果通过量化模型计算得到该节点的威胁值。资产检测器和脆弱点检测器属于静态检测器。这种检测器不会主动检测和匹配抗原，根据免疫网络中与其相邻的动态检测器给予的刺激信息，进行遗传变异，完成其生命周期中的各项活动。根据 2 种检测器的不同功能，在系统中采用了不同的定义方法。

#### 3.2.1　威胁行为检测器的定义

定义信息系统中正常活动为自体集合 $S$，非法活动为非自体集合 $N$，将信息系统中所有的行为数学抽象为一个长度为 $L$ 的二进制字符串集合 $U$，则有 $S \cup N = U$，$S \cap N = \phi$。

定义抗原集合 $Ag$ 为威胁行为的集合( $Ag \subset N$ )，抗体集合为识别威胁行为检测器集合 $Dt$，$Dt$ 也是长度为 $L$ 的二进制字符串集合。当抗体和抗原的亲和力 $d$ 达到某一阈值 $\varepsilon$ 时，视为匹配，即抗体检测出抗原，如式(2)所示。本文亲和力的计算方法建立在 Hamming 距离公式的基础上，具体见式(3)：

$$f_{匹配}(ag,dt)=\begin{cases}1 & d_{亲和力}(ag,dt)>\varepsilon \\ 0 & \text{otherwise}\end{cases} \tag{2}$$

$$d_{亲和力}(ag,dt)=L-\sum_{i=1}^{L}\delta_i \quad \delta_i=\begin{cases}1 & dt_i \neq ag_i \\ 0 & \text{otherwise}\end{cases} \tag{3}$$

其中，$ag \in Ag, dt \in Dt$ ；$dt_i$表示抗体串$dt$的第$i$位；$ag_i$表示抗原串$ag$的第$i$位。

随机生成初始的检测器集合$Dt_0$，对于新产生的检测器，需要经过自体耐受过程才能成熟，如果在耐受过程中与任意自体发生匹配，就将死亡。

耐受过程的数学定义如下：

$$g_{耐受}(x,y)=\left\{ x \in Dt_0, \forall y \in S, f_{区配}(x,y)=0 \right\} \tag{4}$$

经过自体耐受以后，成熟检测器进入免疫循环。如果成熟检测器的亲和力(与抗原匹配次数)达到一定的阈值 $\beta$，就将激活成为记忆检测器，并克隆增殖，以抵御更多抗原的入侵。每个成熟检测器都有固定的生命周期( $\lambda$ )，在周期内未积累足够亲和力的检测器被淘汰。这种新旧更替机制保证了检测器的多样性及其搜索抗原空间的能力。

相对于成熟检测器，记忆检测器的生命周期更长，甚至可以趋于无限长。但是记忆检测器的数量在长时间内是稳定且有限的(模拟人体免疫机制)，当记忆检测器数量达到上限时，可以选择 LRU 等合适的算法进行淘汰。本文选择淘汰最近一段时间内匹配次数最少的记忆检测器。淘汰后的记忆检测器将作为成熟检测器，重新累积亲和力。记忆检测器如果匹配到抗原，将被再次激活，发生克隆增殖。一部分符合条件的记忆检测器还能进行变异，使系统具有进化的能力，变异得到的检测器加入到新生的检测器集合中，需要经过自体耐受过程才能成熟。

#### 3.2.2　静态检测器的定义

每个静态检测器对应一个实数向量，即一组参数。 资产检测器的向量形式为 ***Dan(****aNum, aSet*)，脆弱点检测器的向量形式为 ***Dvk***(*vNum, vSet*)，其中，*aNum* 为该检测器对应资产的编号；*aSet* 为资产重要性，0<*aSet*<1；*vNum* 为脆弱点编号；

*vSet* 为脆弱性程度，0<*vSet*<1。

初始时，静态检测器的个数和参数值根据信息系统的资产清单和漏洞扫描报告生成。威胁行为检测器检测到未知的资产或脆弱点，也会生成相应的资产检测器和脆弱性检测器加入集合中。

当威胁行为检测器检测到威胁，并提取到相关资产编号和脆弱点编号，就会在相应检测器集合中进行编号的匹配，匹配成功则进行联通，形成免疫网络。同时，威胁行为检测器向联通的检测器发出刺激信号，使被刺激的检测器发生遗传变异，并成为记忆检测器。记忆检测器拥有更长的生命周期，*aSet* 值和 *vSet* 值也相应增加。这是因为根据信息系统的规律，发生过的威胁行为再次发生的可能性更高。

静态检测器也有固定的生命周期。在无外界刺激的情况下，*aSet* 值和 *vSet* 值逐渐衰减，当衰减到一定阈值或生命周期结束，该检测器死亡。死亡机制避免了静态检测器集合的无限扩增，及时删除安全节点避免了大量的匹配计算，节省了系统开支。

3.2.3　基于免疫网络的信息安全风险评估学习算法

信息安全风险评估免疫网络模拟了抗原入侵免疫系统及免疫系统产生应答的主要过程。由于抗原的入侵激发威胁行为检测器抗体群种的免疫反应，因此在克隆分化后会引起记忆单元的动态变化并进一步产生免疫克隆增值。威胁行为检测器集合与静态检测器集合构成免疫网络，并激发静态检测器的遗传免疫和记忆过程，完成网络的动态免疫调节。定义$\xi$为选择成熟细胞的比率，$\sigma_{ds}$为相对自然死亡或衰减的阈值，算法如下：

Step1 初始化免疫网络。根据设定，产生抗体群 *Da, Dt, Dv*。

Step2 无抗原或抗原浓度低则停止，否则继续。

Step3 计算抗原和抗体群 *Dt* 中每个抗体的亲和度，选择 *n* 个亲和度高的抗体。

Step4 调节免疫网络。释放原有免疫网络连接，计算 *Dt* 抗体和 *Da, Dv* 中每个抗体的亲和度，建立连接。

Step5 克隆 *Dt* 抗体，使它们的后代数与亲和度成正比。

Step6 亲和度成熟。利用基因操作(主要是变异)增加这部分抗体与抗原的亲和度。*Da, Dv* 中被选择的抗体根据刺激信息进行变异，增加与刺激愿抗体的亲和度。

Step7 更新记忆单元。从 *Dt* 抗体及其后代中选择 $\xi$ 个进入记忆单元 *Mt*，计算 *Mt* 中的亲和度，去除那些亲和度低于 *σds* 的记忆单元，*Da, Dv* 中被选择的抗体进入记忆单元 *Ma, Mv*。

Step8 记忆单元中抗体个数达上限则停止，否则继续。

Step9 记忆单元 *Mt* 亲和度成熟。利用克隆、基因操作(主要是变异)和克隆选择增加记忆单元与抗原的亲和度。

Step10 更新抗体种群。利用抗体及其后代和记忆单元构造新的抗体群。

Step11 返回 Step2。

在 Step6~Step9 中，每个抗体细胞变异程度与父细胞的适应度有关，并根据以下公式进行突变：

$$C' = C + aN(0,1) \tag{5}$$

$$a = (\frac{1}{\beta}e^{-f^*}) \tag{6}$$

其中，细胞 $C'$是细胞 $C$ 变异后产生的新细胞；$N(0,1)$是一个均值为 0、标准偏差为 1 的 Gauss 随机变量；$\beta$ 是用于控制指数函数衰减的变量；$f^*$是经过标准化处理后的细胞适应值(函数取值在[0,1]之间)。

### 3.3　风险计算

安全风险评估模块实时监控免疫网络的拓扑结构，根据各检测器的参数及网络连接情况，计算风险值，方法如下：

(1)设资产 $i$ 在 $t$ 时刻所受到的攻击种类为 $m$，检测器数量分别是 $I_1, I_2, \cdots, I_m$，该资产相对于每类威胁的威胁指数为

$$T_i(t,m) = 1 - \frac{1}{e^{I_m}} \tag{7}$$

(2)资产 $i$ 的风险值为

$$R(i,m) = L_i \times F_i \tag{8}$$

其中，$L_i$ 是资产 $i$ 发生风险事件 $m$ 的可能性；$L_i = \frac{1}{t'}\int_{t'}(T_i(t,m) \times vSet(t))\mathrm{d}t$；$F_i$ 是资产 $i$ 发生风险事件 $m$ 的影响值：$F_i = \frac{1}{t'}\int_{t'}(vSet(t) \times aSet(t))\mathrm{d}t$。根据免疫网络的拓扑结构生成以资产为索引的安全风险报告，包括各资产的风险值，易发生的威胁行为，威胁行为的影响等项。并可根据资产的重要性程度，综合评定该信息系统的风险值。

## 4　实验

选取实验室局域网内的 2 台机器作为目标系统。设置实验参数如下：抗体和抗原的长度均设为 104 位，其中，源地址 32 位；源端口 16 位；目的地址 32 位；目的端口 16 位；协议 2 位；标志位 6 位。成熟检测器的激活阈值 $\beta$ 和生命周期长度 $\lambda$ 分别设为 10 000 s 和 90 000 s。假设受威胁资产的 *aSe*t 分别为 0.4, 0.2，可利用的漏洞脆弱性程度 *vSet* 均为 0.8。生成初始检测器集合。实验中同时对主机 A, B 进行 SYNFLOOD 攻击，对主机 B 进行 Land 攻击，这 2 种攻击的危险指数设为 0.8, 0.6。根据系统的实时监控，得到的结果如图 4、图 5 所示：图 4 描述了攻击强度随时间的变化曲线，图 5 则描绘了系统风险指数随时间变化的曲线，表 1 是系统内各资产的风险分析。

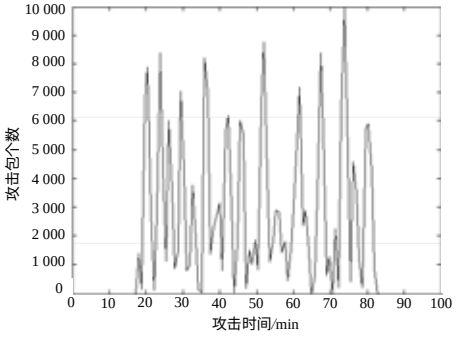


**图 4　攻击强度变化曲线**

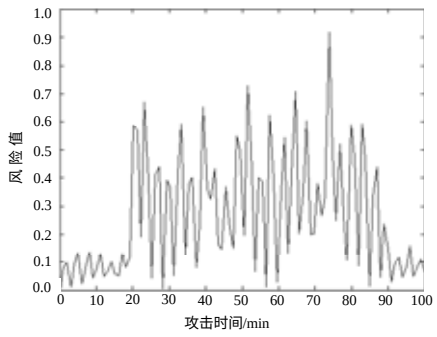


**图 5　遭受攻击时风险指数变化曲线**

（下转第 186 页）